

2005

Average distances in Euclidean and other spaces

Benjamin Michael-John Chad
University of Wollongong

Follow this and additional works at: <https://ro.uow.edu.au/theses>

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation

Chad, Benjamin Michael-John, Average distances in Euclidean and other spaces, Master of Science thesis, School of Mathematics and Applied Statistics, University of Wollongong, 2005.
<https://ro.uow.edu.au/theses/2713>

Average Distances in Euclidean and Other Spaces

*A thesis submitted in partial fulfilment of the requirements
for the award of the degree*

MASTER OF SCIENCE - RESEARCH

from

UNIVERSITY OF WOLLONGONG

by

**BENJAMIN MICHAEL-JOHN CHAD
BACHELOR OF MATHEMATICS (HONOURS)
BACHELOR OF COMPUTER SCIENCE**

**SCHOOL OF MATHEMATICS AND APPLIED STATISTICS .
2005**

I, Benjamin Michael-John Chad, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Master of Science - Research, in the School of Mathematics and Applied Statistics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. This thesis has not been submitted for qualifications at any other academic institution.

Benjamin Michael-John Chad,
December 2005.

Contents

Acknowledgements	iv
Abstract	v
Outline	vii
List of symbols	ix
1 Introduction	1
1.1 The Gross-Stadje Theorem	1
1.2 Suprema of average distances	3
1.2.1 The average distance $M(X, d)$	3
1.2.2 $M(X)$ as a weighted average distance	6
1.2.3 The average distance $\overline{M}(X, d)$	8
1.3 Characterisations of $m(X)$, $M(X)$ and $\overline{M}(X)$	14
1.4 Summary	15
2 Spaces of measures	17
2.1 The space of measures	17
2.2 The strong topology on the space of measures	20
2.3 The weak-* topology on the space of measures	20
2.3.1 Borel measures on a metric space	23
2.3.2 Properties of the weak-* topology on $\mathcal{M}(X)$	27
2.4 Summary	41

3	Average distance integrals	43
3.1	The function $d_\mu: X \rightarrow \mathbb{R}$	43
3.1.1	Properties of d_μ	44
3.2	The functional $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$	46
3.2.1	Continuity properties of $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$	46
3.2.2	Discontinuity of $I: \mathcal{M}([0, 1])^2 \rightarrow \mathbb{R}$	53
3.3	Summary	59
4	Results in compact spaces	61
4.1	The Gross-Stadje Theorem	61
4.2	Properties of $m(X, d)$	68
4.3	Properties of $M(X, d)$	71
4.4	Properties of $\overline{M}(X, d)$	78
4.5	Techniques for calculating $m(X)$ and $M(X)$	80
4.5.1	Max-min measures	81
4.5.2	d -invariant measures	86
4.5.3	The support of max-min and M -maximal measures	93
4.6	Summary	96
5	Quasihypermetric spaces	97
5.1	Hypermetric spaces	97
5.1.1	Characterisations of hypermetric spaces	98
5.1.2	Examples of hypermetric spaces	102
5.2	Quasihypermetric spaces	108
5.2.1	Relationships with hypermetric spaces	109
5.2.2	A characterisation of quasihypermetric spaces	110
5.3	Compact quasihypermetric spaces	111
5.3.1	Measure theoretic characterisations	111
5.3.2	Metric embedding characterisations	114
5.4	Strictly quasihypermetric spaces	130

5.5	Summary	133
6	Results in quasihypermetric spaces	135
6.1	Finiteness of $\overline{M}(X)$	135
6.2	Properties of $m(X)$, $M(X)$ and $\overline{M}(X)$	140
6.2.1	M -maximal and \overline{M} -maximal measures	141
6.2.2	Relationships between $m(X)$, $M(X)$ and $\overline{M}(X)$	142
6.2.3	Discontinuity of $\overline{M}: H(S) \rightarrow \mathbb{R}$	145
6.3	Geometrical interpretations of $M(X)$ and $\overline{M}(X)$	145
6.4	Some values of $m(X)$, $M(X)$ and $\overline{M}(X)$	154
6.5	Summary	156
7	Other work	159
7.1	Averaging powers of distances	159
7.2	Bounds for $m(X)$, $M(X)$ and $\overline{M}(X)$	160
7.3	Examples	161
	Bibliography	163

Acknowledgements

First and foremost, I thank my supervisors Associate Professor Peter Nickolas and Associate Professor Graham Williams. They have both spent substantial time and energy helping me with this project, and they have taught me a considerable amount of mathematics over its duration. I cannot speak highly enough of each of them. In particular, I thank them for their unspoken patience with my progress, and I thank Peter especially for his role as my “primary” supervisor.

I thank the Faculty of Informatics for their financial assistance to give me time to undertake this degree, and I thank Graham Williams for lobbying to obtain this support. It goes without saying that this funding was essential for me to complete this thesis. I also thank the School of Mathematics and Applied Statistics for giving me the opportunity to attend conferences in Melbourne and Auckland.

The submission of this thesis marks, for now, the end of my official association with the University. Reflecting on my time at Wollongong, I am left with mixed emotions, as I will miss my colleagues in the School, but I also look forward to the experiences that await me in the United Kingdom.

It is very hard for someone to undertake a university degree without the support of their family, and this degree has resulted from the support of many people in mine. I thank my best-friend and partner-in-crime Claire for her unfailing support in the face of all the frustration that I cause her, and for bullying me into completing this thesis. I thank our parents for their kind words and encouragement. Finally, I thank Holly Chad for keeping me company on late nights calculating average distances, long after all other mortals have gone to sleep.

Abstract

This thesis is a study of average distances in compact metric spaces. The average distances that we study can be motivated by questions such as “How can one arrange a finite set of points on a square so as to maximise the average distance between them?” The answer to this question is simple: Place a point on each vertex of the square! Such questions become much more interesting, and their corresponding answers become much harder to find, when we generalise the space to other metric spaces, and vary what we mean by the term “average distance”. For instance, is it always possible given any finite set of points on the boundary of a circle to find another point in the same region such that the “average distance” from the point to the set is $2/\pi$? If so, will any other “magic-number” other than $2/\pi$ work?

We define three average distances for a compact metric space (the first as a result of the Gross-Stadje Theorem, which additionally requires that the space be connected, the others as suprema of certain average distances) which form our focus of study. These averages have previously been discussed in mathematical journals, and one purpose of this thesis is to survey much of the known material regarding them. Indeed, until now there has been no single account which discusses all three averages together, and we gain much further insight by doing so.

Consideration of some well-known measure theory quickly changes the nature of our average distances from being considered as properties of finite sets of points, to being considered as properties of (signed) measures of unit mass, for it is possible to characterise our average distances using certain functionals whose domain is a space of measures. We give new results due to Peter Nickolas concerning the (lack

of) continuity of these functionals.

We spend considerable time discussing hypermetric and quasihypermetric spaces. These spaces have previously been studied in mathematical journals under several different names and different characterisations. Until now, there has been no single account which discusses all of these characterisations together. We give new results due to Peter Nickolas and Reinhard Wolf concerning average distances in compact quasihypermetric spaces. Two of our average distances have strong geometric interpretations for certain quasihypermetric spaces, and we present original proofs for metric embedding theorems of Menger which enable us to arrive at these.

Using our measure theoretic framework, we investigate properties of our average distances, and demonstrate how to calculate them for certain concrete spaces, in particular for Euclidean, and more generally, quasihypermetric spaces.

We conclude our discussion by noting certain other questions which we may have chosen to consider in this thesis.

Outline

This thesis is intended to be read in a linear fashion, with the exceptions of Chapter 2 and Chapter 5. The reader who is familiar with weak-* convergence of finite signed Borel measures on a compact metric space, and in particular, convergence in its convex subset of probability measures, may skip Chapter 2, or refer to it as required. Chapter 5 introduces quasihypermetric spaces, and it may be read independently of the other work in this thesis.

Chapter 1 defines the average distances $m(X)$, $M(X)$ and $\overline{M}(X)$ for a compact (connected) metric space X , and calculates each in the case that $X = [0, 1]$. It concludes by noting that these constants may be characterised by integrating the metric of the space with respect to measures of unit mass, and that Chapter 2 and Chapter 3 will develop the theory necessary to do so.

Chapter 2 introduces the topological vector space $\mathcal{M}(X)$ of finite signed Borel measures on a metric space X , and its convex subset $\mathcal{M}^1(X)$ of probability measures, equipped with its weak-* topology. Particular attention is paid to studying certain analytic and topological properties of $\mathcal{M}^1(X)$ in the case that X is compact.

Chapter 3 introduces two mappings derived from integration with respect to signed measures, and discusses properties of each, with particular attention paid to continuity.

Chapter 4 uses the work in the preceding two chapters to prove the Gross-Stadje theorem, to characterise $m(X)$, $M(X)$ and $\overline{M}(X)$ measures, to develop properties of these average distances, and to calculate their values for certain spaces. The discussions in this chapter do not refer to the quasihypermetric inequality.

Chapter 5 introduces the class of quasihypermetric spaces, and the subclasses of hypermetric and strictly quasihypermetric spaces. The choice of results to include in this chapter is motivated by their need for a study of average distances; however, it can be read independently of the remainder of this thesis.

Chapter 6 continues the discussion from Chapter 4, presenting further results on average distances in quasihypermetric spaces.

Chapter 7 notes other questions arising from a consideration of average distances in compact metric spaces which we may have chosen to consider.

List of symbols

$m(X)$	pg. 1,67	The Gross-Stadje number of the compact connected space X .
$D(X)$	pg. 3	The diameter of X .
$M(X)$	pg. 3,7,72	$\sup(1/n^2) \sum_{i=1}^n d(x_i, x_j) = \sup \sum_{i,j=1}^n w_i w_j d(x_i, x_j) = \sup I(\mu)$, where $n \in \mathbb{N}$, $x_i \in X$, $w_i > 0$ such that $\sum_{i=1}^n w_i = 1$ and $\mu \in \mathcal{M}^1(X)$.
$\overline{M}(X)$	pg. 9,78	$\sup \sum_{i,j=1}^n w_i w_j d(x_i, x_j) = \sup I(\mu)$, where $n \in \mathbb{N}$, $x_i \in X$, $w_i \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 1$, and $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$.
$\mathcal{M}(X)$	pg. 14,18,24	The set of finite signed Borel measures on X , equipped with its weak-* topology.
$\mathcal{M}^1(X)$	pg. 14,18,24	The set of finite Borel probability measures on X , equipped with its weak-* topology.
δ_x	pg. 14,19	The atomic Borel probability measure concentrated at $x \in X$.
$\mathcal{M}^+(X)$	pg. 18,24	The set of finite (positive) Borel measures on X , equipped with its weak-* topology.
$\mathcal{B}(X)$	pg. 18	The Borel σ -algebra of X .
$\mu = \mu^+ - \mu^-$	pg. 18	The Jordan decomposition of μ as the difference of two positive measures.
$ \mu $	pg. 18	The variation of μ .
$\ \mu\ $	pg. 18	The total variation of μ .

Chapter 1

Introduction

This thesis is an investigation of average distances in metric spaces, where several different notions of the term “average” are considered. Given a compact metric space X , this chapter introduces the constants $m(X)$, $M(X)$ and $\overline{M}(X)$, each of which in some sense represent an average distance amongst the points of X . We define these constants, briefly mention some relationships between them, and calculate each when X is the unit interval. Finally, we discuss alternate characterisations of $m(X)$, $M(X)$ and $\overline{M}(X)$.

1.1 The Gross-Stadje Theorem

We begin the story of average distances in compact metric spaces with the following surprising result:

Theorem 1.1.1 (Gross-Stadje Theorem). *Let (X, d) be a compact connected metric space. Then there exists a unique $m(X, d) \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ and for all $x_1, \dots, x_n \in X$ there exists $y \in X$ such that*

$$\frac{1}{n} \sum_{i=1}^n d(x_i, y) = m(X, d) .$$

We refer to $m(X, d)$ as the *Gross-Stadje number* of a compact connected metric space (X, d) , and abbreviate $m(X, d)$ to $m(X)$ where appropriate.

The theorem was originally published by Gross [24]. It was also published by Stadje [48], who was apparently unaware of Gross' work. The Stadje result is more general, using the weaker assumptions that is X a compact connected Hausdorff space and that $d: X \times X \rightarrow \mathbb{R}$ is a continuous symmetric function. This thesis is concerned only with the original result of Gross, but we attribute it to both Gross and Stadje.

Gross' proof of the theorem uses game theory, whilst Stadje's proof uses game theory and measure theory. Our proof requires game theory and measure theory, but is different from the one given by Stadje. To allow for a discussion of the appropriate measure theory, we postpone the proof until Chapter 4.

The following proposition specialises the Gross-Stadje Theorem to the compact connected interval $[0, 1]$, equipped with its usual metric. The proof is due to Morris and Nickolas [38, pg. 461].

Proposition 1.1.2. *There exists a unique $m([0, 1]) \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ and for all $x_1, \dots, x_n \in [0, 1]$ there exists $y \in [0, 1]$ such that*

$$\frac{1}{n} \sum_{i=1}^n |x_i - y| = m([0, 1]) .$$

Further, $m([0, 1]) = 1/2$.

Proof. Let $n = 2$, let $x_1 = 0$ and let $x_2 = 1$. Then for each $y \in [0, 1]$,

$$\frac{1}{n} \sum_{i=1}^n |x_i - y| = \frac{1}{2}(|0 - y| + |1 - y|) = \frac{1}{2} .$$

Therefore, if $m([0, 1])$ exists then it must be unique and equal $1/2$.

Let $n \in \mathbb{N}$, let $x_1, \dots, x_n \in [0, 1]$ and let $f: [0, 1] \rightarrow [0, 1]$ be the function defined for $y \in [0, 1]$ by

$$f(y) = \frac{1}{n} \sum_{i=1}^n |x_i - y| .$$

Then f is continuous and

$$f(1) = \frac{1}{n} \sum_{i=1}^n |x_i - 1| = \frac{1}{n} \sum_{i=1}^n (1 - x_i) = 1 - \frac{1}{n} \sum_{i=1}^n x_i = 1 - f(0) .$$

It is easily shown that $f(0) \leq 1/2 \leq f(1)$ or $f(1) \leq 1/2 \leq f(0)$. In either case, by the Intermediate Value Theorem there exists $y \in [0, 1]$ such that $f(y) = 1/2$, thus establishing the existence of $m([0, 1])$. \square

1.2 Suprema of average distances

If X is a compact connected metric space, then $m(X)$ can be considered in some sense to be an average distance which may be attained by using any tuple of points from X . Using different notions of average distance, we now define the constants $M(X)$ and $\overline{M}(X)$ for a compact metric space X . These values represent suprema of mutual average distances in tuples of points, the averages being measured in subtly different ways.

1.2.1 The average distance $M(X, d)$

Let (X, d) be a compact metric space and let $D(X)$ denote the diameter of X . As X is compact, $D(X)$ is a finite non-negative real number and $D(X) = 0$ if and only if X is a singleton. If $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$, then it is clear that

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j) \leq D(X) < \infty.$$

The supremum of average distances of the above form must be finite.

Definition 1.2.1. Let (X, d) be a compact metric space. Then $M(X, d) \in \mathbb{R}$ is defined by

$$M(X, d) = \sup \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j),$$

where the supremum is taken over all $n \in \mathbb{N}$ and all $x_1, \dots, x_n \in X$. The abbreviation $M(X)$ may be used for $M(X, d)$ as appropriate.

When X is also connected, it is not immediately clear that there exists a relationship between the values $m(X)$ and $M(X)$. This relationship will be explored in Chapters 4 and 6. In particular, Chapter 4 will show that $m(X) \leq M(X)$.

For now, we calculate $M([0, 1])$, the derivation of which is due to Peter Nickolas.

Lemma 1.2.2. *Let $x_1, \dots, x_n \in [0, 1]$ contain at least one interior point of $[0, 1]$. Then there exist $y_1, \dots, y_n \in [0, 1]$ and $k \in \{1, \dots, n\}$ such that*

$$\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \leq \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|$$

and $y_i = x_i$ for all i such that $i \neq k$ and x_k is an interior point of $[0, 1]$ and y_k is an endpoint of $[0, 1]$.

Proof. We may assume without loss of generality that the x_i are ordered. As the x_i contain an interior point of $[0, 1]$, let $l = \min \{i: x_i \neq 0\}$ and let $r = \max \{i: x_i \neq 1\}$. Then $l - 1$ is the number of x_i that lie on the left endpoint of $[0, 1]$ and $n - r$ is the number of x_i that lie on the right endpoint of $[0, 1]$. The y_i will be constructed from the x_i by “sliding” some x_j to an endpoint of $[0, 1]$. We consider separately the cases $l - 1 \leq n - r$ and $l - 1 > n - r$.

Case 1: $l - 1 \leq n - r$. We have that the number of x_i which lie on the left endpoint of $[0, 1]$ is less than or equal to the number of x_i which lie on the right endpoint of $[0, 1]$. Hence, let $y_1, \dots, y_n \in [0, 1]$ be constructed from the x_i by sliding x_l to the left endpoint of $[0, 1]$; that is, for each $i = 1, \dots, n$, let

$$y_i = \begin{cases} x_i & \text{if } i \neq l, \\ 0 & \text{if } i = l. \end{cases}$$

As $l - 1 - n + r \leq 0$ and $x_l > 0$, it follows that $(l - 1 - n + r)x_l \leq 0 = (l - 1 - n + r)y_l$. Let $x'_1, \dots, x'_{n-1} \in [0, 1]$ be the list of points created by removing x_l from the x_i , and let $y'_1, \dots, y'_{n-1} \in [0, 1]$ be the list of points created by removing y_l from the y_i . Then as

$$\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} |x'_i - x'_j| = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} |y'_i - y'_j|$$

and

$$\begin{aligned}
\sum_{i=1}^{n-1} |x_l - x'_i| &= \sum_{i=1}^{l-1} |x_l - x_i| + \sum_{i=l+1}^r |x_l - x_i| + \sum_{i=r+1}^n |x_l - x_i| \\
&= (l-1)x_l + \sum_{i=l+1}^r |x_l - x_i| + (n-r)(1-x_l) \\
&= (l-1-n+r)x_l + n-r + \sum_{i=l+1}^r |x_l - x_i| \\
&\leq (l-1-n+r)y_l + n-r + \sum_{i=l+1}^r |y_l - y_i| \\
&= \sum_{i=1}^{n-1} |y_l - y'_i|,
\end{aligned}$$

it follows that

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| &= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} |x'_i - x'_j| + 2 \sum_{i=1}^{n-1} |x_l - x'_i| + |x_l - x_l| \\
&\leq \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} |y'_i - y'_j| + 2 \sum_{i=1}^{n-1} |y_l - y'_i| + |y_l - y_l| \\
&= \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|.
\end{aligned}$$

Case 2: $l-1 > n-r$. We have that the number of x_i which lie on the left endpoint of $[0, 1]$ is greater than the number of x_i which lie on the right endpoint of $[0, 1]$. Hence, let $y_1, \dots, y_n \in [0, 1]$ be constructed from the x_i by sliding x_r to the right endpoint of $[0, 1]$; that is, for each $i = 1, \dots, n$, let

$$y_i = \begin{cases} x_i & \text{if } i \neq r, \\ 1 & \text{if } i = r. \end{cases}$$

As $l-1-n+r > 0$ and $x_r < 1$, it follows that

$$(l-1-n+r)x_r < (l-1-n+r) = (l-1-n+r)y_r.$$

It can be shown in a similar manner to case 1 that

$$\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \leq \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|.$$

□

Proposition 1.2.3. $M([0, 1]) = 1/2$.

Proof. Let $n = 2$, let $x_1 = 0$ and let $x_2 = 1$. Then

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| = \frac{1}{2},$$

which gives that $M([0, 1]) \geq 1/2$. It remains to be shown that $M([0, 1]) \leq 1/2$.

Let $n \in \mathbb{N}$ and let $x_1, \dots, x_n \in [0, 1]$. By applying Lemma 1.2.2 a finite number of times, there exist $y_1, \dots, y_n \in \{0, 1\}$ such that

$$\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \leq \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|.$$

Let k be the cardinality of the set $\{i : 1 \leq i \leq n \text{ and } y_i = 0\}$. Then $n - k$ is the cardinality of the set $\{i : 1 \leq i \leq n \text{ and } y_i = 1\}$ and

$$\sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| = 2k(n - k).$$

Now, $2k(n - k)$ may be considered to be a quadratic in k which attains an upper bound when $k = n/2$. We have that

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \leq \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j| \leq \frac{1}{n^2} \cdot 2 \cdot \frac{n}{2} \cdot \left(n - \frac{n}{2}\right) = \frac{1}{2},$$

giving that $M([0, 1]) \leq 1/2$. □

1.2.2 $M(X)$ as a weighted average distance

We now show that $M(X)$ may be characterised as the supremum of weighted average distances of the form

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the $w_i \geq 0$ are *weights* such that $\sum_{i=1}^n w_i = 1$.

Lemma 1.2.4. *Let $n \in \mathbb{N}$ and let $w_1, \dots, w_n \in [0, 1]$ such that $\sum_{i=1}^n w_i = 1$. Then for all $\varepsilon > 0$ there exist $q_1, \dots, q_n \in \mathbb{Q} \cap [0, 1]$ such that $\sum_{i=1}^n q_i = 1$ and $|w_i - q_i| < \varepsilon$ for each i .*

Proof. Let $\varepsilon > 0$. Then for all i such that $1 \leq i \leq n-1$ there exists $q_i \in \mathbb{Q} \cap [0, 1]$ such that $|w_i - q_i| < \varepsilon/n < \varepsilon$ and $q_i \leq w_i$. Subsequently, let $q_n = 1 - \sum_{i=1}^{n-1} q_i \in \mathbb{Q}$, which gives that $\sum_{i=1}^n q_i = 1$ and

$$0 \leq w_n = 1 - \sum_{i=1}^{n-1} w_i \leq 1 - \sum_{i=1}^{n-1} q_i \leq 1.$$

Hence $q_n \in \mathbb{Q} \cap [0, 1]$, and

$$|w_n - q_n| = \left| \sum_{i=1}^{n-1} (-w_i + q_i) \right| \leq \sum_{i=1}^{n-1} |w_i - q_i| < (n-1) \cdot \frac{\varepsilon}{n} < \varepsilon. \quad \square$$

Proposition 1.2.5. *Let (X, d) be a compact metric space. Then*

$$M(X, d) = \sup \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the supremum is taken over all $n \in \mathbb{N}$ and all $x_1, \dots, x_n \in X$ and all $w_1, \dots, w_n \geq 0$ such that $\sum_{i=1}^n w_i = 1$.

Proof. Let

$$A = \sup \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j),$$

where the supremum is taken over all $n \in \mathbb{N}$ and all $x_1, \dots, x_n \in X$, and let

$$B = \sup \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the supremum is taken over all $n \in \mathbb{N}$ and all $x_1, \dots, x_n \in X$ and all $w_1, \dots, w_n \geq 0$ such that $\sum_{i=1}^n w_i = 1$. We want to show that $A = B$. Noting that $\sum_{i=1}^n (1/n) = 1$, it is clear that $A \leq B$. Let $n \in \mathbb{N}$, let $x_1, \dots, x_n \in X$ and let $w_1, \dots, w_n \geq 0$ such that $\sum_{i=1}^n w_i = 1$. To demonstrate that $B \leq A$, it will be sufficient to show that

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq A,$$

which we shall establish in two steps.

Suppose that each weight is rational. Then there exist $p_1, \dots, p_n \in \mathbb{N} \cup \{0\}$ and $q_1, \dots, q_n \in \mathbb{N}$ such that $w_i = p_i/q_i$ for each i . Let m be the lowest common multiple

of the q_i . We then have that $p_i m / q_i \in \mathbb{Z}$ for each i and

$$\sum_{i=1}^n \frac{p_i m}{q_i} = m \sum_{i=1}^n \frac{p_i}{q_i} = m.$$

Now, let $y_1, \dots, y_m \in X$ be a list of points which contains exactly $p_i m / q_i$ occurrences of x_i for each i . Then

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) = \frac{1}{m^2} \sum_{i=1}^n \sum_{j=1}^n \frac{p_i m}{q_i} \frac{p_j m}{q_j} d(x_i, x_j) = \frac{1}{m^2} \sum_{i=1}^m \sum_{j=1}^m d(y_i, y_j) \leq A.$$

Otherwise there exists an irrational weight. We may assume without loss of generality that the x_i are distinct, and therefore must contain at least two points. Let $\varepsilon > 0$ and let $M = \max \{d(x_i, x_j) : i, j = 1, \dots, n\} > 0$. By Lemma 1.2.4 there exist $q_1, \dots, q_n \in \mathbb{Q} \cap [0, 1]$ such that $\sum_{i=1}^n q_i = 1$ and $|w_i - q_i| < \varepsilon / 2Mn^2$ for each i . Hence

$$\begin{aligned} & \left| \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) - \sum_{i=1}^n \sum_{j=1}^n q_i q_j d(x_i, x_j) \right| \\ & \leq \left| \sum_{i=1}^n \sum_{j=1}^n (w_i w_j - w_i q_j) d(x_i, x_j) \right| + \left| \sum_{i=1}^n \sum_{j=1}^n (w_i q_j - q_i q_j) d(x_i, x_j) \right| \\ & \leq \sum_{i=1}^n \sum_{j=1}^n w_i |w_j - q_j| d(x_i, x_j) + \sum_{i=1}^n \sum_{j=1}^n q_j |w_i - q_i| d(x_i, x_j) \\ & < 2 \sum_{i=1}^n \sum_{j=1}^n 1 \cdot \frac{\varepsilon}{2Mn^2} \cdot M \\ & = \varepsilon, \end{aligned}$$

giving that

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) < \sum_{i=1}^n \sum_{j=1}^n q_i q_j d(x_i, x_j) + \varepsilon \leq A + \varepsilon.$$

By taking the infimum over ε , we obtain $\sum_{i,j=1}^n w_i w_j d(x_i, x_j) \leq A$. □

1.2.3 The average distance $\overline{M}(X, d)$

We now define $\overline{M}(X)$ as a generalisation of $M(X)$ by allowing the possibility of negative weights.

Definition 1.2.6. Let (X, d) be a compact metric space. Then $\overline{M}(X, d) \in \mathbb{R} \cup \{\infty\}$ is defined by

$$\overline{M}(X, d) = \sup \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the supremum is taken over all $n \in \mathbb{N}$ and all $x_1, \dots, x_n \in X$ and all $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 1$. The abbreviation $\overline{M}(X)$ may be used for $\overline{M}(X, d)$ as appropriate.

It is clear that $M(X) \leq \overline{M}(X)$. We note that, unlike $M(X)$, there is no obvious reason why $\overline{M}(X)$ need be finite. In fact, Chapter 6 will give examples of compact metric spaces X such that $\overline{M}(X) = \infty$.

For now, we calculate $\overline{M}([0, 1])$. The following result establishes the so-called *quasihypermetric property* of $[0, 1]$ using a proof due to Kézdy et al. [30, pg. 25-26] (the cited proof actually establishes the stronger *hypermetric property* of $[0, 1]$). We shall study this property further in Chapter 5.

Theorem 1.2.7. For all $n \in \mathbb{N}$ and for all $a_1, \dots, a_n, b_1, \dots, b_n \in [0, 1]$,

$$\sum_{1 \leq i < j \leq n} |a_i - a_j| + \sum_{1 \leq i < j \leq n} |b_i - b_j| \leq \sum_{1 \leq i, j \leq n} |a_i - b_j|.$$

Proof. Let $n \in \mathbb{N}$ and let $a_1, \dots, a_n, b_1, \dots, b_n \in [0, 1]$. Any rearrangement of the a_i clearly leaves $\sum_{i,j=1}^n |a_i - b_j|$ and $\sum_{i,j=1}^n |a_i - a_j|$ invariant, and since

$$\sum_{1 \leq i, j \leq n} |a_i - a_j| = 2 \sum_{1 \leq i < j \leq n} |a_i - a_j|,$$

any rearrangement of the a_i will also leave $\sum_{i < j} |a_i - a_j|$ invariant. Hence, we may assume without loss of generality that the a_i are ordered. Similarly, we may assume without loss of generality that the b_i are ordered.

Let x_1, \dots, x_{2n} be an ordered listing of the points $a_1, \dots, a_n, b_1, \dots, b_n$, and let $a: \{a_1, \dots, a_n\} \rightarrow \{1, \dots, 2n\}$ and $b: \{b_1, \dots, b_n\} \rightarrow \{1, \dots, 2n\}$ be the mappings such that for each i , $x_{a(i)}$ represents the point a_i and $x_{b(i)}$ represents the point b_i . Note that if $i < j$ then $|x_i - x_j| = \sum_{k=i}^{j-1} |x_k - x_{k+1}|$. In general, we may

write

$$|x_i - x_j| = \sum_{k=\min(i,j)}^{\max(i,j)-1} |x_k - x_{k+1}|.$$

Let

$$\begin{aligned} A &= \sum_{1 \leq i < j \leq n} |a_i - a_j| = \sum_{1 \leq i < j \leq n} |x_{a(i)} - x_{a(j)}| = \sum_{1 \leq i < j \leq n} \left(\sum_{k=a(i)}^{a(j)-1} |x_k - x_{k+1}| \right), \\ B &= \sum_{1 \leq i < j \leq n} |b_i - b_j| = \sum_{1 \leq i < j \leq n} \left(\sum_{k=b(i)}^{b(j)-1} |x_k - x_{k+1}| \right), \\ C &= \sum_{1 \leq i, j \leq n} |a_i - b_j| = \sum_{1 \leq i, j \leq n} \left(\sum_{k=\min(a(i), b(j))}^{\max(a(i), b(i))-1} |x_k - x_{k+1}| \right). \end{aligned}$$

For each k , let α_k, β_k and γ_k be the number of formal occurrences of the expression $|x_k - x_{k+1}|$ in each of the expressions defining A, B and C respectively. It will be sufficient to show that for each k , $\alpha_k + \beta_k \leq \gamma_k$.

For each k , let

$$\begin{aligned} T_k &= \{i: i \leq k \text{ and } i = a(j) \text{ for some } 1 \leq j \leq n\}, \\ U_k &= \{i: i \leq k \text{ and } i = b(j) \text{ for some } 1 \leq j \leq n\}, \\ Y_k &= \{i: i > k \text{ and } i = a(j) \text{ for some } 1 \leq j \leq n\}, \\ Z_k &= \{i: i > k \text{ and } i = b(j) \text{ for some } 1 \leq j \leq n\}, \end{aligned}$$

and let t_k, u_k, y_k and z_k be the respective cardinalities of these sets. Note that $t_k + y_k = u_k + z_k = n$. Now, the expression $|x_k - x_{k+1}|$ occurs exactly once in the expression defining A for each pair (i, j) such that $[x_k, x_{k+1}] \subseteq [x_{a(i)}, x_{a(j)}] = [a_i, a_j]$. This is given by the number of pairs (i, j) such that $a(i) \leq k$ and $a(j) \geq k+1 > k$, which is exactly $t_k y_k$. Therefore $\alpha_k = t_k y_k$, and similarly $\beta_k = u_k z_k$. The expression $|x_k - x_{k+1}|$ occurs exactly once in the expression defining C for each pair (i, j) such that $[x_k, x_{k+1}] \subseteq [x_{a(i)}, x_{b(j)}] = [a_i, b_j]$ or $[x_k, x_{k+1}] \subseteq [x_{b(j)}, x_{a(i)}] = [b_j, a_i]$. This is given by the number of pairs (i, j) such that $a(i) \leq k$ and $b(j) > k$, or $b(j) \leq k$ and $a(i) > k$, which is exactly $t_k z_k + u_k y_k$. Therefore $\gamma_k = t_k z_k + u_k y_k$.

For all k , it follows from

$$(u_k - t_k) - (y_k - z_k) = (u_k + z_k) - (t_k + y_k) = n - n = 0$$

that $(u_k - t_k)(y_k - z_k) \geq 0$. Hence $t_k y_k + u_k z_k \leq t_k z_k + u_k y_k$, giving $\alpha_k + \beta_k \leq \gamma_k$ as required. \square

An alternate proof of the fact that \mathbb{R} , and so $[0, 1]$, has the so-called *hypermetric property* is given by Stolarsky [50, pg. 547–549]. The above property of $[0, 1]$ can be stated as a property of weighted average distances. The proof is suggested by Lovász et al. [35, pg. 2049].

Lemma 1.2.8. *Let $n \in \mathbb{N}$ and let $x_1, \dots, x_n \in \mathbb{R}$ such that $\sum_{i=1}^n x_i = 0$. Then for all $\varepsilon > 0$ there exist $q_1, \dots, q_n \in \mathbb{Q}$ such that $\sum_{i=1}^n q_i = 0$ and $|x_i - q_i| < \varepsilon$ for each i .*

Proof. Let $\varepsilon > 0$. Then for all i such that $1 \leq i \leq n-1$ there exists $q_i \in \mathbb{Q}$ such that $|x_i - q_i| < \varepsilon/n < \varepsilon$. Subsequently, let $q_n = -\sum_{i=1}^{n-1} q_i \in \mathbb{Q}$, which gives that $\sum_{i=1}^n q_i = 0$ and

$$|x_n - q_n| = \left| \sum_{i=1}^{n-1} (-x_i + q_i) \right| \leq \sum_{i=1}^{n-1} |x_i - q_i| < (n-1) \cdot \frac{\varepsilon}{n} < \varepsilon. \quad \square$$

Theorem 1.2.9. *For all $n \in \mathbb{N}$, for all $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 0$ and for all $x_1, \dots, x_n \in [0, 1]$,*

$$\sum_{1 \leq i < j \leq n} w_i w_j |x_i - x_j| \leq 0. \quad .$$

Proof. Let $n \in \mathbb{N}$, let $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum w_i = 0$ and let $x_1, \dots, x_n \in [0, 1]$. We shall establish the result in three steps.

Suppose that each weight is an integer. If all of the weights are non-negative then we must have that $w_i = 0$ for each i , giving $\sum_{i < j} w_i w_j |x_i - x_j| = 0$. Otherwise, there exist at least one positive and one negative weight. Let $m \in \mathbb{N}$ be the number of negative weights, let $a_1, \dots, a_m \in [0, 1]$ represent the x_i such that $w_i < 0$ and let $u_1, \dots, u_m \in \mathbb{Z}$ be the corresponding weights, and let $b_1, \dots, b_{n-m} \in [0, 1]$ represent

the x_i such that $w_i \geq 0$ and let $v_1, \dots, v_{n-m} \in \mathbb{Z}$ be the corresponding weights. We then have that

$$\begin{aligned}
& \sum_{1 \leq i < j \leq n} w_i w_j |x_i - x_j| \\
&= \sum_{1 \leq i < j \leq m} u_i u_j |a_i - a_j| + \sum_{1 \leq i < j \leq n-m} v_i v_j |b_i - b_j| + \sum_{i=1}^m \sum_{j=1}^{n-m} u_i v_j |a_i - b_j| \\
&= \sum_{1 \leq i < j \leq m} |u_i| |u_j| |a_i - a_j| + \sum_{1 \leq i < j \leq n-m} |v_i| |v_j| |b_i - b_j| - \sum_{i=1}^m \sum_{j=1}^{n-m} |u_i| |v_j| |a_i - b_j|.
\end{aligned}$$

Let $m' = \sum_{i=1}^m -u_i$ and $n' = \sum_{i=1}^{n-m} v_i$. Since $-m' + n' = 0$, it follows that $m' = n'$. Let $a'_1, \dots, a'_{m'} \in [0, 1]$ be the list of points which contains exactly $|u_i|$ occurrences of a_i for each i and let $b'_1, \dots, b'_{n'} \in [0, 1]$ be the list of points which contains exactly $|v_i|$ occurrences of b_i for each i . Then by Theorem 1.2.7,

$$\begin{aligned}
& \sum_{1 \leq i < j \leq n} |u_i| |u_j| |a_i - a_j| + \sum_{1 \leq i < j \leq n-m} |v_i| |v_j| |b_i - b_j| - \sum_{i=1}^m \sum_{j=1}^{n-m} |u_i| |v_j| |a_i - b_j| \\
&= \sum_{1 \leq i < j \leq m'} |a'_i - a'_j| + \sum_{1 \leq i < j \leq n'} |b'_i - b'_j| - \sum_{1 \leq i, j \leq m'} |a'_i - b'_j| \\
&\leq 0,
\end{aligned}$$

which gives that $\sum_{i < j} w_i w_j |x_i - x_j| \leq 0$.

Suppose now that each weight is rational. Then there exist $p_1, \dots, p_n \in \mathbb{Z}$ and $q_1, \dots, q_n \in \mathbb{N}$ such that $w_i = p_i/q_i$ for each i . Let q be the lowest common multiple of the q_i . It follows that $p_i q/q_i \in \mathbb{Z}$ for each i , and as

$$\sum_{i=1}^n \frac{p_i q}{q_i} = q \sum_{i=1}^n \frac{p_i}{q_i} = 0,$$

we then have

$$\sum_{1 \leq i < j \leq n} w_i w_j |x_i - x_j| = \frac{1}{q^2} \sum_{1 \leq i < j \leq n} \frac{p_i q}{q_i} \frac{p_j q}{q_j} |x_i - x_j| \leq 0.$$

Finally, suppose now that there exists an irrational weight, from which it follows that there must exist at least two distinct points amongst the x_i . Let $\varepsilon > 0$, let

$M = \max \{|x_i - x_j| : i, j = 1, \dots, n\} > 0$, let $K = \max \{2|w_i| : i = 1, \dots, n\} > 0$ and let

$$\delta = \min \left\{ \frac{\varepsilon}{KMn(n-1)}, \frac{K}{2} \right\}.$$

By Lemma 1.2.8, there exist $q_1, \dots, q_n \in \mathbb{Q}$ such that $\sum_{i=1}^n q_i = 0$ and $|w_i - q_i| < \delta$ for each i . Now, for each i , by the triangle inequality we have that

$$||w_i| - |q_i|| = ||w_i - 0| - |q_i - 0|| \leq |w_i - q_i| < \delta$$

giving

$$|q_i| < |w_i| + \delta \leq \max_j |w_j| + \delta \leq \frac{K}{2} + \frac{K}{2} = K.$$

Note that there are exactly $n(n-1)/2$ pairs (i, j) such that $1 \leq i < j \leq n$. Therefore

$$\begin{aligned} & \left| \sum_{1 \leq i < j \leq n} w_i w_j |x_i - x_j| - \sum_{1 \leq i < j \leq n} q_i q_j |x_i - x_j| \right| \\ & \leq \left| \sum_{1 \leq i < j \leq n} (w_i w_j - w_i q_j) |x_i - x_j| \right| + \left| \sum_{1 \leq i < j \leq n} (w_i q_j - q_i q_j) |x_i - x_j| \right| \\ & \leq \sum_{1 \leq i < j \leq n} |w_i| |w_j - q_j| M + \sum_{1 \leq i < j \leq n} |q_j| |w_i - q_i| M \\ & < 2 \sum_{1 \leq i < j \leq n} K \delta M \\ & < 2 \sum_{1 \leq i < j \leq n} \frac{\varepsilon}{n(n-1)} \\ & = \varepsilon, \end{aligned}$$

giving that

$$\sum_{1 \leq i < j \leq n} w_i w_j |x_i - x_j| < \sum_{1 \leq i < j \leq n} q_i q_j |x_i - x_j| + \varepsilon \leq 0 + \varepsilon.$$

By taking the infimum over ε , we obtain $\sum_{i < j} w_i w_j |x_i - x_j| \leq 0$. □

Proposition 1.2.10. $\overline{M}([0, 1]) = 1/2$.

Proof. It is a consequence of Proposition 1.2.3 that $\overline{M}([0, 1]) \geq M([0, 1]) = 1/2$. It remains to be shown that $\overline{M}([0, 1]) \leq 1/2$.

Let $n \in \mathbb{N}$, let $x_1, \dots, x_n \in [0, 1]$ and let $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 1$, let $w_{n+1} = w_{n+2} = -1/2$, let $x_{n+1} = 0$ and let $x_{n+2} = 1$. Then $\sum_{i=1}^{n+2} w_i = 0$ and by Theorem 1.2.7,

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n w_i w_j |x_i - x_j| + \sum_{i=n+1}^{n+2} \sum_{j=n+1}^{n+2} w_i w_j |x_i - x_j| + 2 \sum_{i=1}^n \sum_{j=n+1}^{n+2} w_i w_j |x_i - x_j| \\ &= \sum_{i=1}^{n+2} \sum_{j=1}^{n+2} w_i w_j |x_i - x_j| \\ &\leq 0. \end{aligned}$$

Now,

$$\sum_{i=n+1}^{n+2} \sum_{j=n+1}^{n+2} w_i w_j |x_i - x_j| = \frac{1}{4} (0 + 1 + 1 + 0) = \frac{1}{2}.$$

Noting that $|x_i - 0| + |x_i - 1| = 1$ for all $x \in [0, 1]$, we then have

$$2 \sum_{i=1}^n \sum_{j=n+1}^{n+2} w_i w_j |x_i - x_j| = - \sum_{i=1}^n (w_i |x_i - 0| + w_i |x_i - 1|) = - \sum_{i=1}^n w_i = -1.$$

Hence

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j |x_i - x_j| \leq 1/2,$$

which gives that $\overline{M}([0, 1]) \leq 1/2$. □

1.3 Characterisations of $m(X)$, $M(X)$ and $\overline{M}(X)$

Let (X, d) be a compact metric space. Let $\mathcal{M}(X)$ denote the set of finite signed Borel measures on X and let $\mathcal{M}^1(X)$ denote the subset of $\mathcal{M}(X)$ consisting of precisely the probability measures. For each $x \in X$, let δ_x denote the atomic Borel probability measure concentrated at x . Observe that for each atomic measure $\mu = \sum_{i=1}^n w_i \delta_{x_i} \in \mathcal{M}(X)$ and for all $y \in X$,

$$\begin{aligned} \int d(x, y) d\mu(x) &= \sum_{i=1}^n w_i d(x_i, y), \\ \int \int d(x, z) d\mu(x) d\mu(z) &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j). \end{aligned}$$

The constants $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$ may then be characterised using various integrals of the metric d with respect to atomic measures.

Now, consider

$$M(X, d) = \sup_{\mu} \int d(x, y) d\mu(x) d\mu(y).$$

We initially defined $M(X, d)$, in effect, by taking the supremum over atomic measures of the form $\mu = (1/n) \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}^1(X)$. Proposition 1.2.5 showed that this value is not altered by allowing the μ to range over an enlarged class of measures, namely the collection of finitely supported Borel probability measures on X . Chapter 4 will give a measure theoretic characterisation of $M(X, d)$ by further enlarging this class of measures to include all Borel probability measures on X . By replacing the probability measures with signed measures of unit mass, analogous comments apply to $\overline{M}(X, d)$.

If (X, d) is connected then $m(X, d) \in \mathbb{R}$ is the unique constant such that for all atomic measures of the form $\mu = (1/n) \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}^1(X)$ there exists $y \in X$ such that

$$\int d(x, y) d\mu(x) = m(X, d).$$

Chapter 4 will prove a strong form of the Gross-Stadje Theorem which characterises $m(X, d)$ using arbitrary Borel probability measures on X .

To characterise $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$ using measures, and to prove the Gross-Stadje Theorem, we must first discuss some measure theory. Chapter 2 will study two topologies which may be placed on $\mathcal{M}(X)$, and Chapter 3 will study properties of certain mappings concerning the integrals

$$\int d(x, y) d\mu(x) \quad \text{and} \quad \int d(x, z) d\mu(x) d\mu(z),$$

where $y \in X$ and $\mu \in \mathcal{M}(X)$.

1.4 Summary

This chapter stated the Gross-Stadje Theorem and defined the average distances $m(X)$, $M(X)$ and $\overline{M}(X)$ found in a compact (connected) metric space X . From

these definitions, it is clear that $M(X) \leq \overline{M}(X)$, but it is not clear that any further relationship exists between $M(X)$ and $\overline{M}(X)$, nor given the Gross-Stadje Theorem that $m(X)$ is related to these values. Relationships between the three average distances will be developed in Chapters 4 and 6, but it was shown that when $X = [0, 1]$ all three averages are defined and equal to $1/2$.

We noted that $m(X)$, $M(X)$ and $\overline{M}(X)$ may be characterised by integrating various functions with respect to Borel measures of unit mass. To derive these characterisations, and to prove the Gross-Stadje Theorem, it is necessary to discuss convergence in $\mathcal{M}(X)$ and properties of certain functions which arise by integrating the metric of X . These topics are discussed in Chapters 2 and 3. The remainder of this thesis will then be a study of $m(X)$, $M(X)$ and $\overline{M}(X)$, and the relationships between them.

Chapter 2

Spaces of measures

This chapter will introduce the vector space $\mathcal{M}(X)$ of finite signed Borel measures on a metric space X , and discuss the strong and weak-* topologies on this space. Particular attention will be paid to the properties of convergence, compactness, metrization and approximation in $\mathcal{M}(X)$ equipped with the weak-* topology, which will subsequently be used in Chapter 4 to prove the Gross-Stadje Theorem, and to investigate certain properties of $m(X)$, $M(X)$ and $\overline{M}(X)$.

2.1 The space of measures

We include the following definition to fix the meaning of the terms *signed measure*, *measure* and *probability measure*.

Definition 2.1.1. Let (X, \mathcal{X}) be a measurable space and let $\mu: \mathcal{X} \rightarrow [-\infty, \infty]$ be a function. If $\mu(\emptyset) = 0$ and

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

for all disjoint sequences $(A_n) \in \mathcal{X}$ then μ is a *signed measure*. In addition, if $\mu(A) \geq 0$ for all $A \in \mathcal{X}$ then μ is a *measure*. Further, if μ is a measure such that $\mu(X) = 1$ then μ is a *probability measure*.

According to Bartle [6, pg. 23], as the left-hand side of the above equality is independent of order and required for all disjoint sequences (A_n) , the series on the right-hand side must be absolutely convergent for all such sequences. In particular, this requires that μ may only attain one of the values ∞ and $-\infty$.

Our measure theory notation will follow that of Cohn [17]. In particular, we shall employ his convention of not abbreviating the term *signed measure* to *measure*, though for emphasis we will at times refer to a *measure* as a *positive measure*. If a (signed) measure μ does not take on the values $\pm\infty$, then we say that μ is a *finite (signed) measure*.

Let X be a metric space. Then $\mathcal{B}(X)$ will denote the Borel σ -algebra on X ; that is, $\mathcal{B}(X)$ is the σ -algebra generated by the open (or equivalently, closed) subsets of X . We refer to a (signed) measure on the measurable space $(X, \mathcal{B}(X))$ as a *(signed) Borel measure*, and we define

$$\begin{aligned}\mathcal{M}(X) &= \{\mu: \mu \text{ is a finite signed Borel measure on } X\}, \\ \mathcal{M}^+(X) &= \{\mu: \mu \text{ is a finite Borel measure on } X\}, \\ \mathcal{M}^1(X) &= \{\mu: \mu \text{ is a Borel probability measure on } X\}.\end{aligned}$$

It is clear that $\mathcal{M}^1(X) \subseteq \mathcal{M}^+(X) \subseteq \mathcal{M}(X)$. It is well-known that $\mathcal{M}(X)$ is a real vector space under the usual operations of addition and scalar multiplication, and that $\mathcal{M}^+(X)$ and $\mathcal{M}^1(X)$ are convex subsets of $\mathcal{M}(X)$. The subsequent sections of this chapter will discuss various topologies on $\mathcal{M}(X)$.

Let $\mu \in \mathcal{M}(X)$. Then $\mu = \mu^+ - \mu^-$ will denote the unique *Jordan decomposition* of μ as the difference of two positive measures. (See [17, pg. 125] for the standard derivation of μ^+ and μ^- .) The measure $|\mu| = \mu^+ + \mu^-$ is the *variation* of μ , and the real number $\|\mu\| = |\mu|(X)$ is the *total variation* of μ . The total variation is a norm on $\mathcal{M}(X)$, and $\mathcal{M}(X)$ is a Banach space when equipped with this norm. We define the support of μ to be the set

$$C_\mu = \{x \in X: |\mu|(U) > 0 \text{ for all open neighbourhoods } U \text{ of } x\}.$$

We denote by $C(X)$ the Banach space of bounded continuous real-valued func-

tions on X equipped with the usual supremum norm, and by $U(X)$ the subspace of uniformly continuous functions. For $f \in C(X)$ and $\mu \in \mathcal{M}(X)$, we note that the integral of f with respect to μ is defined, and as convenient use the notation

$$\mu(f) = \int f d\mu.$$

We refer to a measure with finite support as an *atomic* measure. For each $x \in X$, the atomic measure concentrated at x , denoted δ_x , is the probability measure such that for all $A \in \mathcal{B}(X)$,

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

The atomic probability measures play an important role in approximation of measures with respect to the weak-* topology in $\mathcal{M}(X)$. For now, we see that these measures are also important to the convex structure of $\mathcal{M}^1(X)$. The following result is given as an exercise in Choquet [14, pg. 108].

Proposition 2.1.2. *Let X be a compact metric space. Then the set of probability measures $\{\delta_x : x \in X\}$ is the set of extreme points of $\mathcal{M}^1(X)$.*

Proof. Let $x \in X$, let $\mu, \nu \in \mathcal{M}^1(X)$ be such that $\delta_x = (\mu + \nu)/2$ and let $A \in \mathcal{B}(X)$. If $x \in A$ then $\mu(A) + \nu(A) = 2$ and as $0 \leq \mu, \nu \leq 1$ it follows that $\mu(A) = \nu(A) = 1$. Otherwise $x \notin A$ and $\mu(A) + \nu(A) = 0$, and as μ, ν are measures it follows that $\mu(A) = \nu(A) = 0$. Hence $\delta_x = \mu = \nu$ and δ_x is an extreme point of $\mathcal{M}^1(X)$. \bullet

Let $\mu \in \mathcal{M}^1(X)$ be such that $\mu \neq \delta_x$ for all $x \in X$. Since X is a compact metric space, the support of μ is measurable. It cannot be empty, since μ is a probability measure, and we have excluded the case that it is a singleton. The support must then contain at least two points, x_1 and x_2 say. Let $U \in \mathcal{B}(X)$ be an open neighbourhood of x_1 such that $x_2 \notin U$, whence $0 < \mu(U), \mu(X \setminus U) < 1$. Let $\nu, \lambda \in \mathcal{M}^1(X)$ be the measures defined for $A \in \mathcal{B}(X)$ by

$$\nu(A) = \frac{1}{\mu(U)}\mu(A \cap U) \quad \text{and} \quad \lambda(A) = \frac{1}{\mu(X \setminus U)}\mu(A \cap (X \setminus U)).$$

Then for all $A \in \mathcal{B}(X)$,

$$\begin{aligned}\mu(A) &= \mu(A \cap U) + \mu(A \cap (X \setminus U)) \\ &= \mu(U) \cdot \nu(A) + (1 - \mu(U)) \cdot \lambda(A),\end{aligned}$$

giving that $\mu = \mu(U)\nu + (1 - \mu(U))\lambda$ is not an extreme point of $\mathcal{M}^1(X)$. \square

2.2 The strong topology on the space of measures

Definition 2.2.1. Let V be a set which is both a real vector space and a Hausdorff topological space, and consider $V \times V$ and $\mathbb{R} \times V$ to be equipped with their respective product topologies. If

1. The mapping $V \times V \rightarrow V$ such that $(x, y) \mapsto x + y$ is continuous, and
2. The mapping $\mathbb{R} \times V \rightarrow V$ such that $(x, y) \mapsto x \cdot y$ is continuous,

then V is a real topological vector space.

If X is a metric space, then we refer to the norm topology on $\mathcal{M}(X)$ as the *strong topology on $\mathcal{M}(X)$* . Using elementary properties of the norm, this space is a real topological vector space. We denote strong convergence of a net $(\mu_\alpha) \in \mathcal{M}(X)$ to a limit $\mu \in \mathcal{M}(X)$ by $\mu_\alpha \rightarrow \mu$. We note that limits of convergent nets in Hausdorff spaces are unique.

2.3 The weak-* topology on the space of measures

Let X be a metric space. We now define another topology on $\mathcal{M}(X)$ by constructing a neighbourhood system about each point. Let $\mu \in \mathcal{M}(X)$. For each $f \in C(X)$ and for each $\varepsilon > 0$, let

$$\dot{V}_\mu(f, \varepsilon) = \left\{ \nu \in \mathcal{M}(X) : \left| \int f d\nu - \int f d\mu \right| < \varepsilon \right\}.$$

For each $n \in \mathbb{N}$, for each $f_1, \dots, f_n \in C(X)$ and for each $\varepsilon_1, \dots, \varepsilon_n > 0$, let

$$V_\mu(f_1, \dots, f_n, \varepsilon_1, \dots, \varepsilon_n) = \bigcap_{i=1}^n V_\mu(f_i, \varepsilon_i).$$

The following two results are stated without proof by Parthasarathy [43, pg. 40].

Proposition 2.3.1. *Let X be a metric space and let \mathcal{B} be the collection of all $V_\mu(f_1, \dots, f_n, \varepsilon_1, \dots, \varepsilon_n)$ such that $\mu \in \mathcal{M}(X)$ and $n \in \mathbb{N}$ and $f_1, \dots, f_n \in C(X)$ and $\varepsilon_1, \dots, \varepsilon_n > 0$. Then \mathcal{B} is a basis for a topology on $\mathcal{M}(X)$.*

Proof. Let $\mu \in \mathcal{M}(X)$ and let 0 denote the zero of $C(X)$. Then for all $\nu \in \mathcal{M}(X)$,

$$\left| \int 0 d\nu - \int 0 d\mu \right| = 0 < 1.$$

Therefore $\mathcal{M}(X) = V_\mu(0, 1) \in \mathcal{B}$.

Let $B_1, B_2 \in \mathcal{B}$ such that $B_1 \cap B_2 \neq \emptyset$. Then $B_1 = V_\mu(f_1, \dots, f_m, \varepsilon_1, \dots, \varepsilon_m)$ and $B_2 = V_\nu(g_1, \dots, g_n, \delta_1, \dots, \delta_n)$ for some $\mu, \nu \in \mathcal{M}(X)$, for some $m, n \in \mathbb{N}$, for some $f_1, \dots, f_m, g_1, \dots, g_n \in C(X)$ and for some $\varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n > 0$. If $\mu = \nu$ then

$$B_1 \cap B_2 = V_\mu(f_1, \dots, f_m, g_1, \dots, g_n, \varepsilon_1, \dots, \varepsilon_m, \delta_1, \dots, \delta_n) \in \mathcal{B}.$$

Otherwise, let $\lambda \in B_1 \cap B_2$. For each $i = 1, \dots, m$ and for each $j = 1, \dots, n$ we have that $\varepsilon_i - |\lambda(f_i) - \mu(f_i)| > 0$ and $\delta_j - |\lambda(g_j) - \nu(g_j)| > 0$. Let

$$r = \min \left\{ \varepsilon_i - \left| \int f_i d\lambda - \int f_i d\mu \right|, \delta_j - \left| \int g_j d\lambda - \int g_j d\nu \right| \right\} > 0.$$

Let $V_\lambda = V_\lambda(f_1, \dots, f_m, g_1, \dots, g_n, r, \dots, r, r, \dots, r) \in \mathcal{B}$. It is clear that $\lambda \in V_\lambda$. Let $\omega \in V_\lambda$. Then for each $i = 1, \dots, m$

$$\begin{aligned} \left| \int f_i d\omega - \int f_i d\mu \right| &\leq \left| \int f_i d\omega - \int f_i d\lambda \right| + \left| \int f_i d\lambda - \int f_i d\mu \right| \\ &< r + \left| \int f_i d\lambda - \int f_i d\mu \right| \\ &\leq \varepsilon_i - \left| \int f_i d\lambda - \int f_i d\mu \right| + \left| \int f_i d\lambda - \int f_i d\mu \right| \\ &= \varepsilon_i. \end{aligned}$$

That is, $|\omega(f_i) - \mu(f_i)| < \varepsilon_i$ for each $i = 1, \dots, m$. Similarly, $|\omega(g_j) - \nu(g_j)| < \delta_j$ for each $j = 1, \dots, n$, giving $\omega \in B_1 \cap B_2$. We have that $V_\lambda \subseteq B_1 \cap B_2$, and

$$B_1 \cap B_2 = \bigcup_{\lambda \in B_1 \cap B_2} V_\lambda.$$

Therefore \mathcal{B} is a basis for a topology on $\mathcal{M}(X)$. □

We now show that the topology generated by the $V_\mu(f_1, \dots, f_n, \varepsilon_1, \dots, \varepsilon_n)$ is known as the *weak-* topology on $\mathcal{M}(X)$* . This topology is often studied within a general framework in functional analysis.

If X is a Banach space then the dual space of X , denoted by X^* , is the Banach space of all bounded linear functionals on X .

Definition 2.3.2. Let X be a Banach space and let X^* be the dual space of X . A topology on X^* such that a net $(\lambda_\alpha) \in X^*$ converges to $\lambda \in X^*$ if and only if $\lambda_\alpha(x)$ converges to $\lambda(x)$ for all $x \in X$ is known as a *weak-* topology on X^** .

A weak-* topology on a space is necessarily unique. Now, and in later work, we require the following well-known theorem of Riesz. Note that $C(X)$ is a Banach space for all metric spaces X .

Theorem 2.3.3 (Riesz Representation Theorem). *Let X be a metric space. Then for all $\lambda \in C(X)^*$, there exists a unique $\mu \in \mathcal{M}(X)$ such that for all $f \in C(X)$,*

$$\lambda(f) = \int f d\mu \quad \text{and} \quad \|\lambda\| = \|\mu\|.$$

Further, if λ is non-negative and $\lambda(1) = 1$ then $\mu \in \mathcal{M}^1(X)$.

It is now possible to characterise our topology on the space of measures.

Proposition 2.3.4. *Let X be a metric space. Then the topology on $\mathcal{M}(X)$ generated by the $V_\mu(f_1, \dots, f_n, \varepsilon_1, \dots, \varepsilon_n)$ is its weak-* topology.*

Proof. Recall that for all $f \in C(X)$ and for all $\mu \in \mathcal{M}(X)$, it is convenient to write

$$\mu(f) = \int f d\mu.$$

It is a consequence of the Riesz Representation Theorem that the space $C(X)^*$ with its usual norm may be represented by the space of measures $\mathcal{M}(X)$ with the usual measure norm. Let $(\mu_\alpha) \in \mathcal{M}(X)$ be a net and let $\mu \in \mathcal{M}(X)$. We want to show that (μ_α) converges to μ in the given topology if and only if for all $f \in C(X)$, $\mu_\alpha(f) \rightarrow \mu(f)$.

Suppose that μ_α converges to μ in the given topology. Let $f \in C(X)$ and let $\varepsilon > 0$. As $V_\mu(f, \varepsilon)$ is an open neighbourhood of μ there exists an α_0 such that for all α

$$\alpha_0 \leq \alpha \implies \mu_\alpha \in V_\mu(f, \varepsilon) \implies \left| \int f d\mu_\alpha - \int f d\mu \right| < \varepsilon.$$

Therefore $\mu_\alpha(f) \rightarrow \mu(f)$.

Conversely, suppose that for all $f \in C(X)$, $\mu_\alpha(f) \rightarrow \mu(f)$. Let $g_1, \dots, g_n \in C(X)$ and let $\varepsilon_1, \dots, \varepsilon_n > 0$. Then $V_\mu(g_1, \dots, g_n, \varepsilon_1, \dots, \varepsilon_n)$ is an open neighbourhood of μ . Let $\varepsilon_0 = \min \{\varepsilon_i : 1 \leq i \leq n\} > 0$. For each i , there exists α_i such that for all α

$$\alpha_i \leq \alpha \implies \left| \int g_i d\mu_\alpha - \int g_i d\mu \right| < \varepsilon_0.$$

Let α_0 be an upper bound of the α_i . Then for all α and for each i ,

$$\alpha_0 \leq \alpha \implies \left| \int g_i d\mu_\alpha - \int g_i d\mu \right| < \varepsilon_0 < \varepsilon_i.$$

That is, $\alpha_0 \leq \alpha \implies \mu_\alpha \in V_\mu(g_1, \dots, g_n, \varepsilon_1, \dots, \varepsilon_n)$. Therefore μ_α converges to μ in the given topology. \square

Unless stated otherwise, we assume from now on that $\mathcal{M}(X)$, $\mathcal{M}^+(X)$ and $\mathcal{M}^1(X)$ are topological spaces equipped with the weak-* topology. Convergence of a net $(\mu_\alpha) \in \mathcal{M}(X)$ to a limit $\mu \in \mathcal{M}(X)$ in this space will be denoted by $\mu_\alpha \rightarrow^* \mu$.

The remainder of this section will investigate the properties of convergence, compactness, metrization and approximation in $\mathcal{M}(X)$ and $\mathcal{M}^1(X)$. We first derive some necessary measure theoretic results

2.3.1 Borel measures on a metric space

Let (X, d) be a metric space. For all subsets A of X , the mapping $X \rightarrow \mathbb{R}$ such that $x \mapsto d(x, A)$ is defined for $x \in X$ by $d(x, A) = \inf \{d(x, a) : a \in A\}$.

Proofs of the following two lemmas are found in Parthasarathy [43, pg. 2,4].

Lemma 2.3.5. *Let (X, d) be a metric space and let A be a subset of X . Then the mapping $X \rightarrow \mathbb{R}$ such that $x \mapsto d(x, A)$ is uniformly continuous.*

Proof. Let $x, y \in X$. Then for all $z \in A$, $d(x, A) \leq d(x, z) \leq d(x, y) + d(y, z)$ and by taking the infimum over $z \in A$, we obtain $d(x, A) \leq d(x, y) + d(y, A)$. Similarly, $d(y, A) \leq d(x, y) + d(x, A)$ gives that $|d(x, A) - d(y, A)| \leq d(x, y)$. The result follows. \square

The following is a special case of Urysohn's Lemma for normal topological spaces.

Lemma 2.3.6. *Let (X, d) be a metric space and let A and B be disjoint closed subsets of X . Then there exists $f \in C(X)$ such that $0 \leq f \leq 1$ and $f(A) = \{0\}$ and $f(B) = \{1\}$. In addition, if $\inf \{d(a, b) : a \in A \text{ and } b \in B\} > 0$ then f may be chosen such that $f \in U(X)$.*

Proof. Note that as A and B are closed and disjoint, $d(x, A) + d(x, B) \neq 0$ for all $x \in X$. Hence, let $f : X \rightarrow \mathbb{R}$ be the function defined for $x \in X$ by

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}.$$

Then $0 \leq f \leq 1$ and $f(A) = \{0\}$ and $f(B) = \{1\}$, and it is a consequence of Lemma 2.3.5 that $f \in C(X)$. Suppose that $\inf \{d(a, b) : a \in A \text{ and } b \in B\} > 0$ and let $\delta_0 = \inf \{d(x, A) + d(x, B) : x \in X\}$. Then as A and B are closed and disjoint, it follows that $\delta_0 > 0$ and hence for all $x, y \in X$,

$$\begin{aligned} & |f(x) - f(y)| \\ & \leq \left| \frac{d(x, A)}{d(x, A) + d(x, B)} - \frac{d(y, A)}{d(x, A) + d(x, B)} \right| + \left| \frac{d(y, A)}{d(x, A) + d(x, B)} - f(y) \right| \\ & \leq \frac{1}{d(x, A) + d(x, B)} |d(x, A) - d(y, A)| \\ & \quad + \frac{f(y)}{d(x, A) + d(x, B)} \cdot \left| \frac{d(y, A)}{f(y)} - (d(x, A) + d(x, B)) \right| \\ & \leq \frac{1}{\delta_0} \cdot |d(x, A) - d(y, A)| + \frac{1}{\delta_0} |(d(y, A) + d(y, B)) - (d(x, A) + d(x, B))| \\ & \leq \frac{2}{\delta_0} \cdot |d(x, A) - d(y, A)| + \frac{1}{\delta_0} \cdot |d(x, B) - d(y, B)|. \end{aligned}$$

Let $\varepsilon > 0$. As $d(\cdot, A)$ and $d(\cdot, B)$ are uniformly continuous, there exists $\delta > 0$ such that for all $x, y \in X$, if $d(x, y) < \delta$ then $|d(x, A) - d(y, A)| < \delta_0 \varepsilon / 3$ and $|d(x, B) - d(y, B)| < \delta_0 \varepsilon / 3$. We have that for all $x, y \in X$,

$$d(x, y) < \delta \implies |f(x) - f(y)| < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

giving that $f \in U(X)$. □

Let X be a metric space and let $\mu \in \mathcal{M}^1(X)$. Parthasarathy shows in [43, pg. 27] that μ is *regular*; that is, for all $A \in \mathcal{B}(X)$,

$$\begin{aligned} \mu(A) &= \inf \{ \mu(U) : A \subseteq U \text{ is an open subset of } X \} \\ &= \sup \{ \mu(C) : C \subseteq A \text{ is a closed subset of } X \}. \end{aligned}$$

Consequently, μ is uniquely determined by its values on the open or closed subsets of X . The following result shows that a signed measure is uniquely determined by the values it assigns to the integrals of bounded uniformly continuous real-valued functions on X . The proof is an extension of one given by Parthasarathy [43, pg. 39] for probability measures.

Proposition 2.3.7. *Let X be a metric space and let $\mu, \nu \in \mathcal{M}(X)$. If for all $f \in U(X)$,*

$$\int f d\mu = \int f d\nu,$$

then $\mu = \nu$.

Proof. Firstly, suppose that $\mu, \nu \in \mathcal{M}^+(X)$ and that for all $f \in U(X)$,

$$\int f d\mu = \int f d\nu.$$

Let C be a closed subset of X and let (G_n) be a sequence of subsets of X defined for each $n \in \mathbb{N}$ by $G_n = \{x \in X : d(x, C) < 1/n\}$. As $d(\cdot, C)$ is continuous, we have that each G_n is open. Now, (G_n) is a decreasing sequence such that $C = \bigcap_{n=1}^{\infty} G_n$ and for each n , C and $X \setminus G_n$ are disjoint closed sets such that

$$\inf \{d(x, y) : x \in C \text{ and } y \in X \setminus G_n\} \geq 1/n.$$

By Urysohn's Lemma, there exists a sequence (f_n) in $U(X)$ such that for each n , $0 \leq f_n \leq 1$ and $f_n(X \setminus G_n) = \{0\}$ and $f_n(C) = \{1\}$. It is easy seen that for each n , $\chi_C \leq f_n \leq \chi_{G_n}$, giving that

$$\mu(C) = \int \chi_C d\mu \leq \int f_n d\mu = \int f_n d\nu \leq \int \chi_{G_n} d\nu = \nu(G_n).$$

It follows that $\mu(C) \leq \nu(C)$. Similarly $\nu(C) \leq \mu(C)$, and $\mu(C) = \nu(C)$. As μ is regular, it follows that $\mu = \nu$.

Suppose now that $\mu, \nu \in \mathcal{M}(X)$ and that for all $f \in U(X)$,

$$\int f d\mu = \int f d\nu.$$

Let $\mu = \mu^+ - \mu^-$ and $\nu = \nu^+ - \nu^-$ be Jordan decompositions of μ and ν . It is easily seen that $\mu^+(f) + \nu^-(f) = \nu^+(f) + \mu^-(f)$ for all $f \in U(X)$, and from the above, $\mu^+ + \nu^- = \nu^+ + \mu^-$. Therefore $\mu = \nu$. \square

Corollary 2.3.8. *Let X be a metric space and let $\mu \in \mathcal{M}(X)$. If for all $f \in U(X)$*

$$\int f d\mu = 0,$$

then $\mu = 0$.

The proof of the following result was suggested by Peter Nickolas.

Lemma 2.3.9. *Let μ be a finite signed Borel measure on some metric space. Then μ has at most a countable number of points of positive mass.*

Proof. Let A be the set of points given positive mass by μ , which we suppose to be uncountable, and let P be a positive set for μ . Now, for some $m \in \mathbb{N}$, A contains an uncountable number of points with mass greater than $1/m$, otherwise A could be written as a countable union of countable sets. Let $(u_n) \in A$ be a sequence of points with mass greater than $1/m$. Then for each n , $\bigcup_{i=1}^n \{u_i\}$ and its relative

complement in P are μ -measurable and

$$\begin{aligned}
 \mu(A) &= \mu\left(\bigcup_{i=1}^n \{u_i\}\right) + \mu\left(P \setminus \bigcup_{i=1}^n \{u_i\}\right) \\
 &\geq \mu\left(\bigcup_{i=1}^n \{u_i\}\right) \\
 &= \sum_{i=1}^n \mu(\{u_i\}) \\
 &\geq \frac{n}{m}.
 \end{aligned}$$

Letting $n \rightarrow \infty$ we obtain $\mu(A) = \infty$, which contradicts that μ is finite. \square

2.3.2 Properties of the weak-* topology on $\mathcal{M}(X)$

The following result is well-known.

Proposition 2.3.10. *Let X be a metric space. Then $\mathcal{M}(X)$ equipped with its weak-* topology is a real topological vector space.*

Proof. We know that $\mathcal{M}(X)$ is a vector space. Let $(\mu_\alpha, \nu_\alpha) \in \mathcal{M}(X) \times \mathcal{M}(X)$ be a convergent net with limit $(\mu, \nu) \in \mathcal{M}(X) \times \mathcal{M}(X)$. As the product topology on $\mathcal{M}(X) \times \mathcal{M}(X)$ is the weakest topology on this set such that the projection mappings $(\lambda, \omega) \mapsto \lambda$ and $(\lambda, \omega) \mapsto \omega$ are continuous, we must have that $\mu_\alpha \rightarrow^* \mu$ and $\nu_\alpha \rightarrow^* \nu$. Hence, for all $f \in C(X)$

$$\left| \int f d(\mu_\alpha + \nu_\alpha) - \int f d(\mu + \nu) \right| \leq \left| \int f d\mu_\alpha - \int f d\mu \right| + \left| \int f d\nu_\alpha - \int f d\nu \right| \rightarrow 0,$$

giving that $\mu_\alpha + \nu_\alpha \rightarrow^* \mu + \nu$. Therefore, addition is continuous. \bullet

Let $(c_\alpha, \mu_\alpha) \in \mathbb{R} \times \mathcal{M}(X)$ be a convergent net with limit $(c, \mu) \in \mathbb{R} \times \mathcal{M}(X)$. By properties of the product topology on $\mathbb{R} \times \mathcal{M}(X)$, we must have that $c_\alpha \rightarrow c$ and $\mu_\alpha \rightarrow^* \mu$. Hence, for all $f \in C(X)$

$$\begin{aligned}
 \left| \int f d(c_\alpha \mu_\alpha) - \int f d(c\mu) \right| &\leq \left| \int f d(c_\alpha \mu_\alpha) - \int f d(c_\alpha \mu) \right| + \left| \int f d(c_\alpha \mu) - \int f d(c\mu) \right| \\
 &\leq |c_\alpha| \cdot \left| \int f d\mu_\alpha - \int f d\mu \right| + \int \|f\| d|c_\alpha \mu - c\mu| \\
 &= |c_\alpha| \cdot \left| \int f d\mu_\alpha - \int f d\mu \right| + \|f\| \cdot |c_\alpha - c| \cdot \|\mu\|.
 \end{aligned}$$

Let $\varepsilon > 0$. As $c_\alpha \rightarrow c$ there exists α_1 such that

$$\alpha_1 < \alpha \implies |c_\alpha - c| < \frac{\varepsilon}{2(\|f\| \cdot \|\mu\| + 1)}.$$

Also $|c_\alpha| \rightarrow |c|$, giving the existence of an α_2 such that $\alpha_2 < \alpha \implies |c_\alpha| < 2(|c| + 1)$.

As $\mu_\alpha \rightarrow^* \mu$ there exists α_3 such that

$$\alpha_3 < \alpha \implies \left| \int f d\mu_\alpha - \int f d\mu \right| < \frac{\varepsilon}{4(|c| + 1)}.$$

Let α_0 be an upper bound of $\{\alpha_1, \alpha_2, \alpha_3\}$. If $\alpha_0 < \alpha$ then

$$\begin{aligned} \left| \int f d(c_\alpha \mu_\alpha) - \int f d(c\mu) \right| &< 2(|c| + 1) \cdot \frac{\varepsilon}{4(|c| + 1)} + \|f\| \cdot \|\mu\| \cdot \frac{\varepsilon}{2(\|f\| \cdot \|\mu\| + 1)} \\ &= \frac{\varepsilon}{2} + \frac{\|f\| \cdot \|\mu\|}{\|f\| \cdot \|\mu\| + 1} \cdot \frac{\varepsilon}{2} \\ &< \varepsilon, \end{aligned}$$

giving that $c_\alpha \mu_\alpha \rightarrow^* c\mu$. Therefore, scalar multiplication is continuous.

It remains to be shown that $\mathcal{M}(X)$ is Hausdorff. As $(\mathcal{M}(X), +)$ is a topological group, it will be sufficient to show that two distinct points can be separated by an open set. Let $\mu, \nu \in \mathcal{M}(X)$ such that $\mu \neq \nu$. Then by Proposition 2.3.7 there exists $f \in C(X)$ such that $\mu(f) \neq \nu(f)$. Let $r = |\mu(f) - \nu(f)| > 0$. Then $V_\mu(f, r)$ is an open neighbourhood of μ such that $\nu \notin V_\mu(f, r)$. Therefore $\mathcal{M}(X)$ is Hausdorff. \square

The following properties of $\mathcal{M}(X)$ concern convergence.

Proposition 2.3.11. *Let X be a metric space. Then the strong topology on $\mathcal{M}(X)$ contains the weak-* topology on $\mathcal{M}(X)$; that is, for all nets $(\mu_\alpha) \in \mathcal{M}(X)$ and all $\mu \in \mathcal{M}(X)$, if $\mu_\alpha \rightarrow \mu$ then $\mu_\alpha \rightarrow^* \mu$.*

Proof. Suppose that $\mu_\alpha \rightarrow \mu$. Then $\|\mu_\alpha - \mu\| \rightarrow 0$, and for all $f \in C(X)$

$$\left| \int f d\mu_\alpha - \int f d\mu \right| \leq \int \|f\| d|\mu_\alpha - \mu| = \|f\| \cdot \|\mu_\alpha - \mu\| \rightarrow 0.$$

Therefore $\mu_\alpha \rightarrow^* \mu$. \square

Let A be a subset of a topological space X . We denote by A^0 the interior of A , by \bar{A} the closure of A and by $\text{Bd}(A) = \bar{A} - A^0$ the boundary of A .

Lemma 2.3.12. *Let X and Y be topological spaces, let A be a subset of Y and let $f: X \rightarrow Y$ be a continuous function. Then $\text{Bd}(f^{-1}(A)) \subseteq f^{-1}(\text{Bd}(A))$.*

Proof. Note that interior and closure are monotonic functions on subsets of a topological space with respect to subset inclusion and that $A^0 \subseteq A \subseteq \overline{A}$. As f is continuous and $f^{-1}(A^0) \subseteq f^{-1}(A) \subseteq f^{-1}(\overline{A})$, it follows from A^0 being open that $f^{-1}(A^0) \subseteq (f^{-1}(A))^0$, and from \overline{A} being closed that $\overline{f^{-1}(A)} \subseteq f^{-1}(\overline{A})$. Hence

$$\text{Bd}(f^{-1}(A)) = \overline{f^{-1}(A)} - (f^{-1}(A))^0 \subseteq f^{-1}(\overline{A}) - f^{-1}(A^0) = f^{-1}(\text{Bd}(A)). \quad \square$$

The proof of the following result is found in Parthasarathy [43, pg. 40]

Theorem 2.3.13. *Let X be a metric space, let $(\mu_\alpha) \in \mathcal{M}^1(X)$ be a net and let $\mu \in \mathcal{M}^1(X)$. The following are equivalent:*

1. $\mu_\alpha \rightarrow^* \mu$.
2. For all $f \in C(X)$, $\mu_\alpha(f) \rightarrow \mu(f)$.
3. For all $f \in U(X)$, $\mu_\alpha(f) \rightarrow \mu(f)$.
4. For all $C \in \mathcal{B}(X)$, if C is closed then $\limsup_\alpha \mu_\alpha(C) \leq \mu(C)$.
5. For all $U \in \mathcal{B}(X)$, if U is open then $\liminf_\alpha \mu_\alpha(U) \geq \mu(U)$.
6. For all $A \in \mathcal{B}(X)$, if $\mu(\text{Bd}(A)) = 0$ then $\mu_\alpha(A) \rightarrow \mu(A)$. •

Proof. Our usual characterisation of weak-* convergence of probability measures gives that (1) \iff (2). That (2) \implies (3) follows immediately from $U(X) \subseteq C(X)$.

To show that (3) \implies (4), suppose that $\mu_\alpha(f) \rightarrow \mu(f)$ for all $f \in U(X)$. Let C be a closed subset of X and let (U_n) be the sequence of open subsets of X defined for $n \in \mathbb{N}$ by $U_n = \{x \in X : d(x, C) < 1/n\}$. Then (U_n) and $(\mu(U_n))$ are non-increasing sequences such that $C = \bigcap_{n=1}^\infty U_n$ and $\mu(U_n) \rightarrow \mu(C)$. For each n , C and $X \setminus U_n$ are disjoint closed sets such that $\inf \{d(x, y) : x \in C \text{ and } y \in X \setminus U_n\} \geq 1/n$, and by Urysohn's Lemma there exists $f_n \in U(X)$ such that $f_n(X \setminus U_n) = \{0\}$ and

$f_n(C) = \{1\}$ and $0 \leq f_n \leq 1$. Then $\chi_C \leq f_n \leq \chi_{U_n}$ and

$$\begin{aligned} \limsup_{\alpha} \mu_{\alpha}(C) &= \limsup_{\alpha} \int \chi_C d\mu_{\alpha} \\ &\leq \limsup_{\alpha} \int f_n d\mu_{\alpha} \\ &= \int f_n d\mu \\ &\leq \int \chi_{U_n} d\mu \\ &\leq \mu(U_n). \end{aligned}$$

Letting $n \rightarrow \infty$, we obtain $\limsup_{\alpha} \mu_{\alpha}(C) \leq \mu(C)$.

That (4) \iff (5) follows from properties of measures and open and closed sets.

To show that (5) \implies (6), suppose that $\liminf_{\alpha} \mu_{\alpha}(U) \geq \mu(U)$ for all open sets U . Then $\limsup_{\alpha} \mu_{\alpha}(C) \leq \mu(C)$ for all closed sets C . Let $A \in \mathcal{B}(X)$ such that $\mu(\text{Bd}(A)) = 0$. Then $\mu(A^0) = \mu(A) = \mu(\overline{A})$ and

$$\begin{aligned} \limsup_{\alpha} \mu_{\alpha}(A) &\leq \limsup_{\alpha} \mu_{\alpha}(\overline{A}) \leq \mu(\overline{A}) = \mu(A), \\ \liminf_{\alpha} \mu_{\alpha}(A) &\geq \liminf_{\alpha} \mu_{\alpha}(A^0) \geq \mu(A^0) = \mu(A). \end{aligned}$$

Therefore $\mu_{\alpha}(A) \rightarrow \mu(A)$.

To show that (6) \implies (2), suppose that $\mu_{\alpha}(A) \rightarrow \mu(A)$ for all $A \in \mathcal{B}(X)$ such that $\mu(\text{Bd}(A)) = 0$. Let $A \in \mathcal{B}(X)$ such that $\mu(\text{Bd}(A)) = 0$, let $f \in C(X)$ and let $\nu \in \mathcal{M}^1(\mathbb{R})$ be the probability measure defined for $B \in \mathcal{B}(\mathbb{R})$ by $\nu(B) = \mu(f^{-1}(B))$. As f is bounded, there exists $a, b \in \mathbb{R}$ such that $a < b$ and $f(X) \subseteq (a, b)$. Then ν is supported in (a, b) .

Let $\varepsilon > 0$. By Lemma 2.3.9, ν can have at most a countable number of points of non-zero mass. Hence, each set in a finite partition of \mathbb{R} into Borel sets must contain a point that is not a mass point of ν . Then there exist $a = y_0 < y_1 < \dots < y_n = b$ such that $|y_i - y_{i-1}| < \varepsilon/4$ for each $i = 1, \dots, n$ and $\nu(\{y_i\}) = \mu(f^{-1}(\{y_i\})) = 0$ for each $i = 0, \dots, n$.

For each $i = 1, \dots, n$, let $A_i = f^{-1}([y_{i-1}, y_i]) \in \mathcal{B}(X)$. It is a consequence of Lemma 2.3.12 that $\mu(\text{Bd}(A_i)) \leq \mu(f^{-1}(\{y_{i-1}, y_i\})) = 0$, giving $\mu_{\alpha}(A_i) \rightarrow \mu(A_i)$. Let

$g = \sum_{i=1}^n y_i \chi_{A_i}$ be a simple function on X . Then $\|f - g\| < \varepsilon/4$ and

$$\begin{aligned}
& \left| \int f d\mu_\alpha - \int f d\mu \right| \\
& \leq \left| \int (f - g) d\mu_\alpha \right| + \left| \int g d\mu_\alpha - \int g d\mu \right| + \left| \int (g - f) d\mu \right| \\
& \leq \int \|f - g\| d\mu_\alpha + \left| \sum_{i=1}^n y_i \mu_\alpha(A_i) - \sum_{i=1}^n y_i \mu(A_i) \right| + \int \|g - f\| d\mu \\
& < \frac{\varepsilon}{2} + \sum_{i=1}^n |y_i| \cdot |\mu_\alpha(A_i) - \mu(A_i)|.
\end{aligned}$$

As $\mu_\alpha(A_i) \rightarrow \mu(A_i)$ for each i , there exists α_0 such that

$$\alpha_0 \leq \alpha \implies \sum_{i=1}^m |y_i| \cdot |\mu_\alpha(A_i) - \mu(A_i)| < \varepsilon/2.$$

Then $\alpha_0 \leq \alpha \implies |\mu_\alpha(f) - \mu(f)| < \varepsilon$. That is, $\mu_\alpha(f) \rightarrow \mu(f)$. \square

We now investigate metrizable and compactness properties of the spaces of measures $\mathcal{M}(X)$ and $\mathcal{M}^1(X)$. We will make use of the Stone-Weierstrass Theorem, a proof of which may be found in Engelking [22, pg. 144].

Theorem 2.3.14 (The Stone-Weierstrass Theorem). *Let X be a compact metric space, and let $A \subseteq C(X)$ be a ring which contains all constant functions, separates points, and is closed with respect to uniform convergence. Then $A = C(X)$.*

A proof of the following result may be found in Parthasarathy [43, pg. 43].

Proposition 2.3.15. *Let X be a totally bounded metric space. Then $U(X)$ is separable.*

Proof. Let Y be the completion of X . Since Y is complete and totally bounded, it follows that Y is compact. We will now show that $U(X)$ and $C(Y)$ are isometric Banach spaces. Noting that X is dense in Y , for a fixed $f \in U(X)$, let $g_f: Y \rightarrow \mathbb{R}$ be the extension of f defined for $x \in Y \setminus X$ by

$$g_f(x) = \lim_{n \rightarrow \infty} f(x_n),$$

where $(x_n) \in X$ is some sequence converging to x . We need to show that the definition of $g_f(x)$ is independent of the particular sequence chosen to approximate point $x \in Y \setminus X$. Let $(x_n), (y_n) \in X$ be convergent sequences with limit $y \in Y \setminus X$. Then (x_n) and (y_n) are Cauchy sequences, and as f is uniformly continuous, $(f(x_n))$ and $(f(y_n))$ are also Cauchy. By the completeness of Y , there exist $y, z \in Y$ such that $\lim f(x_n) = y$ and $\lim f(y_n) = z$. Given that the sequence $x_1, y_1, x_2, y_2, \dots$ is Cauchy, there exists $w \in Y$ such that $f(x_1), f(y_1), f(x_2), f(y_2), \dots$ has limit w . Then $w = y = z$, as y and z are limits of subsequences of a sequence with limit w . Therefore

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} f(y_n),$$

as required. It is clear that $g_f \in C(Y) = U(Y)$, and so the mapping $U(X) \rightarrow C(Y)$ such that $f \mapsto g_f$ is a bijection. Further,

$$\sup_{x \in X} |f(x)| = \sup_{x \in Y} |g_f(x)|,$$

and so $U(X)$ is isometric to $C(Y)$.

It suffices now to establish that $C(Y)$ is separable. Since Y is compact, it is also second-countable. Let \mathcal{B} be a countable basis for the topology on Y such that $\emptyset \in \mathcal{B}$, and let $\mathcal{C} = \{(U, V) \in \mathcal{B}^2 : \overline{U} \subseteq V\}$. For each $(U, V) \in \mathcal{C}$, by Lemma 2.3.6 there exists $f_{U,V} \in C(Y)$ such that $f(U) \subseteq \{0\}$ and $f(V) \subseteq \{1\}$. Let A be the set of all finite linear combinations of products of functions $f_{U,V}$ with rational coefficients. Then A is a countable ring, and so \overline{A} is a ring. Since Y is regular, it follows that A , and then \overline{A} , separates points. Finally, $f_{\emptyset, X} \equiv 1$ and so \overline{A} contains all constant functions. By the Stone-Weierstrass Theorem, we have that $\overline{A} = C(Y)$, whence $C(Y)$ is separable. \square

Parthasarathy [43, pg. 42-46] supplies proofs of the following three results.

Lemma 2.3.16. *Let X be a metric space. Then X is homeomorphic to the collection of probability measures $\{\delta_x : x \in X\}$.*

Proof. Let $\phi : X \rightarrow \{\delta_x : x \in X\}$ be the function defined for $x \in X$ by $\phi(x) = \delta_x$. It is clear that ϕ is bijective. Let $(x_\alpha) \in X$ be a convergent net with limit $x \in X$.

Then for all $f \in C(X)$,

$$\int f d\delta_{x_\alpha} = f(x_\alpha) \rightarrow f(x) = \int f d\delta_x.$$

Therefore $\phi(x_\alpha) \rightarrow^* \phi(x)$, giving that ϕ is continuous.

Let $(x_\alpha) \in X$ be a net such that $\phi(x_\alpha) \rightarrow^* \phi(x)$ for some $x \in X$. Suppose that $x_\alpha \not\rightarrow x$. There exists an open neighbourhood U of x and a subnet $(x_{\alpha(\beta)})$ of (x_α) such that $x_{\alpha(\beta)} \notin U$ for all β . Note that $\delta_{x_{\alpha(\beta)}} \rightarrow^* \delta_x$. As $\{x\}$ and $X \setminus U$ are disjoint closed sets, by Urysohn's Lemma there exists $f \in C(X)$ such that $0 \leq f \leq 1$ and $f(\{x\}) = \{0\}$ and $f(X \setminus U) = \{1\}$. It follows that $\delta_{x_{\alpha(\beta)}}(f) = 1$ for all β and $\delta_x(f) = 0$. Hence

$$\int f d\delta_{x_{\alpha(\beta)}} \not\rightarrow \int f d\delta_x,$$

contradicting that $\delta_{x_\beta} \rightarrow^* \delta_x$. Therefore $x_\alpha \rightarrow x$ and ϕ^{-1} is continuous. \square

Lemma 2.3.17. *Let X be a metric space. Then $\{\delta_x : x \in X\}$ is a sequentially closed subset of $\mathcal{M}^1(X)$.*

Proof. Let $(x_n) \in X$ be a sequence such that $(\delta_{x_n}) \in \{\delta_x : x \in X\}$ converges to some $\mu \in \mathcal{M}^1(X)$. Suppose that (x_n) does not contain a convergent subsequence. Let $A = \{x_n : n \in \mathbb{N}\}$. Then A must be infinite, otherwise (x_n) would contain a constant subsequence. Let A_1 and A_2 be disjoint infinite sets such that $A = A_1 \cup A_2$. If A_1 has an adherent point y , then by considering the family of open spheres with radius $1/n$ centred at y for $n \in \mathbb{N}$ we can construct a convergent subsequence of (x_n) . Therefore A_1 has no adherent points, and is hence closed. By Theorem 2.3.13,

$$\limsup_n \delta_{x_n}(A_1) \leq \mu(A_1) \leq 1.$$

As A_1 is infinite, for all n there exists m such that $n < m$ and $x_m \in A_1$, giving that

$$\limsup_n \delta_{x_n}(A_1) = 1.$$

It follows that $\mu(A_1) = 1$. Similarly, A_2 is closed and $\mu(A_2) = 1$, giving the contradiction $\mu(A) = 2$. Therefore, (x_n) contains a convergent subsequence (x_{n_k}) . Let $x \in X$ such that $x_{n_k} \rightarrow x$. Then $\delta_{x_{n_k}} \rightarrow^* \delta_x$, giving that $\mu = \delta_x \in \{\delta_x : x \in X\}$. \square

Note that in general, the term *sequentially closed* characterises a weaker condition than the term *closed*, but that these conditions are equivalent in a separable space.

Theorem 2.3.18. *Let X be a metric space. Then*

1. $\mathcal{M}^1(X)$ is separable and metrizable if and only if X is separable, and
2. $\mathcal{M}^1(X)$ is compact and metrizable if and only if X is compact.

Proof. Recall that X is homeomorphic to a sequentially closed subspace of $\mathcal{M}^1(X)$. It is then immediate that if $\mathcal{M}^1(X)$ is separable and metrizable then X is separable, and if $\mathcal{M}^1(X)$ is additionally compact then X is compact.

Suppose then that X is separable. We may equip X with a totally bounded metric, in which case it follows from Proposition 2.3.15 that $U(X)$ is separable. Let $S = \{f \in U(X) : \|f\| \leq 1\}$ and let $\{f_1, f_2, \dots\}$ be a dense subset of S . Then for all n and for all $\mu \in \mathcal{M}^1(X)$, $\mu(f_n) \in [-1, 1]$. Let $\phi : \mathcal{M}^1(X) \rightarrow [-1, 1]^\infty$ be the function defined for $\mu \in \mathcal{M}^1(X)$ by

$$\phi(\mu) = \left(\int f_1 d\mu, \int f_2 d\mu, \dots \right).$$

We want to show that ϕ (with a suitably restricted co-domain) is a homeomorphism.

To show that ϕ is one-to-one, let $\mu, \nu \in \mathcal{M}^1(X)$ such that $\phi(\mu) = \phi(\nu)$. Then $\mu(f_n) = \nu(f_n)$ for all n . To show that $\mu = \nu$, by Proposition 2.3.7 and elementary properties of the integral, it will be sufficient to show that $\mu(f) = \nu(f)$ for all $f \in S$. Let $f \in S$. Then as $\{f_1, f_2, \dots\}$ is dense in S , there exists a subsequence (f_{n_k}) of (f_n) such that $\|f - f_{n_k}\| \rightarrow 0$. Let $\varepsilon > 0$. There exists k_1 such that if $k \geq k_1$ then $\|f - f_{n_k}\| < \varepsilon/2$, giving

$$\left| \int f d\mu - \int f_{n_k} d\mu \right| \leq \int \|f - f_{n_k}\| d\mu = \|f - f_{n_k}\| < \varepsilon/2.$$

Similarly, there exists k_2 such that if $k \geq k_2$ then $|\nu(f) - \nu(f_{n_k})| < \varepsilon/2$. Let k_0 be the maximum of k_1 and k_2 . Then

$$\left| \int f d\mu - \int f d\nu \right| \leq \left| \int f d\mu - \int f_{n_{k_0}} d\mu \right| + \left| \int f_{n_{k_0}} d\nu - \int f d\nu \right| < \varepsilon.$$

Taking the infimum over ε , it follows that $\mu(f) = \nu(f)$ and ϕ is one-one.

To show that ϕ is continuous, let $(\mu_\alpha) \in \mathcal{M}^1(X)$ be a convergent net with limit $\mu \in \mathcal{M}^1(X)$. Then for each n , $\mu_\alpha(f_n) \rightarrow \mu(f_n)$, and by properties of the product topology, $\phi(\mu_\alpha) \rightarrow \phi(\mu)$. Hence ϕ is continuous.

To show that ϕ^{-1} is continuous, let (μ_α) be a net in $\mathcal{M}^1(X)$ and $\mu \in \mathcal{M}^1(X)$ such that $\phi(\mu_\alpha) \rightarrow \phi(\mu)$. Then for all n , $\mu_\alpha(f_n) \rightarrow \mu(f_n)$. We want to show that $\mu_\alpha(f) \rightarrow \mu(f)$ for all $f \in C(X)$. Using Theorem 2.3.13 and properties of the integral, it will be sufficient to prove the statement for $f \in S$. Let $f \in S$. Then for all n and for all α ,

$$\begin{aligned} & \left| \int f d\mu_\alpha - \int f d\mu \right| \\ & \leq \left| \int f d\mu_\alpha - \int f_n d\mu_\alpha \right| + \left| \int f_n d\mu_\alpha - \int f_n d\mu \right| + \left| \int f_n d\mu - \int f d\mu \right| \\ & \leq 2\|f - f_n\| + \left| \int f_n d\mu_\alpha - \int f_n d\mu \right|. \end{aligned}$$

Let $\varepsilon > 0$. There exists n_0 such that $\|f - f_{n_0}\| < \varepsilon/4$. As $\mu_\alpha(f_{n_0}) \rightarrow \mu(f_{n_0})$ there exists α_0 such that if $\alpha_0 \leq \alpha$ then $|\mu_\alpha(f_{n_0}) - \mu(f_{n_0})| < \varepsilon/2$, and so $|\mu_\alpha(f) - \mu(f)| < \varepsilon$. Hence $\mu_\alpha \rightarrow^* \mu$, and so ϕ^{-1} is continuous. We have shown that ϕ (with a suitably restricted co-domain) is a homeomorphism. Recall that the Hilbert cube $[-1, 1]^\infty$ is separable and metrizable. Hence, $\mathcal{M}^1(X)$ is separable and metrizable.

Suppose further that X is compact. As the Hilbert cube is also compact, to show that $\mathcal{M}^1(X)$ is metrizable, it will be sufficient to show that $\phi(\mathcal{M}^1(X))$ is closed. Let $(\mu_n) \in \mathcal{M}^1(X)$ be a sequence such that $(\phi(\mu_n)) \in \phi(\mathcal{M}^1(X))$ is Cauchy, and hence convergent in the Hilbert cube. We want to show that there exists $\mu \in \mathcal{M}^1(X)$ such that $\phi(\mu)$ is the limit point of this sequence.

As $(\phi(\mu_n))$ is convergent, $(\mu_n(f_k))$ is convergent and hence Cauchy for each k .

Let $f \in S$. Then for all m, n and k ,

$$\begin{aligned} & \left| \int f d\mu_m - \int f d\mu_n \right| \\ & \leq \left| \int f d\mu_m - \int f_k d\mu_m \right| + \left| \int f_k d\mu_m - \int f_k d\mu_n \right| + \left| \int f_k d\mu_n - \int f d\mu_n \right| \\ & \leq 2\|f - f_k\| + \left| \int f_k d\mu_m - \int f_k d\mu_n \right|. \end{aligned}$$

By a similar argument to above, $(\mu_n(f))$ is Cauchy and by properties of the integral, $(\mu_n(f))$ is Cauchy for all $f \in U(X) = C(X)$. As a Cauchy sequence of real numbers is convergent, let $\lambda: C(X) \rightarrow \mathbb{R}$ be the linear functional defined for $f \in C(X)$ by

$$\lambda(f) = \lim_{n \rightarrow \infty} \int f d\mu_n.$$

Then λ is a nonnegative linear functional such that $\lambda(1) = 1$. By the Riesz Representation Theorem there exists $\mu \in \mathcal{M}^1(X)$ such that $\lambda(f) = \mu(f)$ for all $f \in C(X)$.

Then for all k

$$\lim_{n \rightarrow \infty} \int f_k d\mu_n = \int f_k d\mu,$$

giving that

$$\begin{aligned} \lim_{n \rightarrow \infty} \phi(\mu_n) &= \left(\lim_{n \rightarrow \infty} \int f_1 d\mu_n, \lim_{n \rightarrow \infty} \int f_2 d\mu_n, \dots \right) \\ &= \left(\int f_1 d\mu, \int f_2 d\mu, \dots \right) \\ &= \phi(\mu). \end{aligned}$$

Hence $\phi(\mathcal{M}^1(X))$ is closed and $\mathcal{M}^1(X)$ is metrizable. □

Corollary 2.3.19. *Let X be a compact metric space. Then $\mathcal{M}^1(X)$ is complete.*

We state the following results of Choquet [13, pg. 217, 308] without proof.

Proposition 2.3.20. *Let X be a compact metric space. Then $\{\mu \in \mathcal{M}(X) : \|\mu\| \leq r\}$ is weak-* compact for $0 \leq r < \infty$.*

Proposition 2.3.21. *Let X be a compact metric space. Then $\mathcal{M}(X)$ is metrizable if and only if X is finite.*

Our remaining results concern approximation in $\mathcal{M}(X)$ and $\mathcal{M}^1(X)$.

Lemma 2.3.22. *Let X be a separable metric space and let $\mu \in \mathcal{M}(X)$ such that its support C_μ is countably infinite. Then $\mu = \sum_{i=1}^{\infty} w_i \delta_{x_i}$ for some sequence $(x_n) \in C_\mu$ and some sequence $(w_n) \in \mathbb{R}$ such that the series $\sum_{i=1}^{\infty} w_i$ is convergent. Further, if $\mu \in \mathcal{M}^+(X)$ then the w_i are non-negative.*

Proof. We note that for each $x \in X$ and each $f \in C(X)$,

$$\int f d\delta_x = f(x).$$

Then by using linearity properties of the integral,

$$\int f d\left(\sum_{i=1}^n w_i \delta_{y_i}\right) = \sum_{i=1}^n w_i f(y_i)$$

for all $n \in \mathbb{N}$ and all $w_1, \dots, w_n \in \mathbb{R}$ and all $y_1, \dots, y_n \in X$. We now list the elements of C_μ as $C_\mu = \{x_1, x_2, \dots\}$. Since X is a separable metric space, X is Lindelöf, and so by the countable subadditivity of a measure, it is easily shown that $|\mu|(\{X \setminus C_\mu\}) = 0$, in which case, $\mu(C_\mu) = \mu(X)$.

We now establish the result in three steps. Firstly, suppose that $\mu \in \mathcal{M}^1(X)$. Let $(w_n) \in [0, 1]$ and $(\mu_n) \in \mathcal{M}^+(X)$ be the sequences defined for $n \in \mathbb{N}$ by $w_n = \mu(\{x_n\})$ and $\mu_n = \sum_{i=1}^n w_i \delta_{x_i}$. We want to show that $\mu_n \rightarrow^* \mu$. Now, $(\mu_n(X))$ is a non-decreasing sequence which is bounded above by 1, and so by the Monotone Convergence Theorem must have a limit. Noting that for each n , $\mu_n(X) = \sum_{i=1}^n w_i$, it follows that $\sum_{i=1}^{\infty} w_i$ is convergent. Let $(A_n) \in \mathcal{B}(X)$ be the sequence defined for $n \in \mathbb{N}$ by $A_n = \bigcup_{i=1}^n \{x_i\}$. Then for each $n \in \mathbb{N}$, we have that the sequence of measurable functions $(\chi_{\{x_{n+1}, \dots, x_k\}})_{k \geq n+1}$ converges pointwise to the measurable function $\chi_{C_\mu \setminus A_n}$, and so by the Lebesgue Dominated Convergence Theorem it follows that

$$\begin{aligned} \mu(C_\mu \setminus A_n) &= \int \chi_{C_\mu \setminus A_n} d\mu \\ &= \lim_{k \rightarrow \infty} \int \chi_{\{x_{n+1}, \dots, x_k\}} d\mu \\ &= \lim_{k \rightarrow \infty} \mu(\{x_{n+1}, \dots, x_k\}) \end{aligned}$$

$$\begin{aligned}
&= \lim_{k \rightarrow \infty} \sum_{i=n+1}^k \mu(\{x_i\}) \\
&= \sum_{i=n+1}^{\infty} \mu(\{x_i\}) \\
&= \sum_{i=n+1}^{\infty} w_i.
\end{aligned}$$

Note that $\sum_{i=n+1}^{\infty} w_i$ must be convergent. Then for all $f \in C(X)$, recalling that $\mu(X) = \mu(C_\mu)$, we have

$$\begin{aligned}
\left| \int f d\mu_n - \int f d\mu \right| &= \left| \sum_{i=1}^n w_i f(x_i) - \left(\sum_{i=1}^n \int_{\{x_i\}} f d\mu + \int_{X \setminus A_n} f d\mu \right) \right| \\
&= \left| \sum_{i=1}^n w_i f(x_i) - \sum_{i=1}^n w_i f(x_i) - \int_{X \setminus A_n} f d\mu \right| \\
&\leq \|f\| \cdot \mu(X \setminus A_n) \\
&= \|f\| \cdot \mu(C_\mu \setminus A_n) \\
&= \|f\| \cdot \sum_{i=n+1}^{\infty} w_i,
\end{aligned}$$

which by the convergence of $\sum_{i=1}^{\infty} w_i$ necessarily tends to 0 as $n \rightarrow \infty$. Hence $\mu_n(f) \rightarrow \mu(f)$ for all $f \in C(X)$, giving that $\mu_n \rightarrow^* \mu$.

Secondly, suppose that $\mu \in \mathcal{M}^+(X)$. As $C_\mu \neq \emptyset$, we must have that $\mu \neq 0$. Given that $(1/\mu(X)) \cdot \mu \in \mathcal{M}^1(X)$, let $(w_n) \in [0, 1]$ and let $(x_n) \in C_\mu$ be sequences such that $(1/\mu(X)) \cdot \mu = \sum_{i=1}^{\infty} w_i \delta_{x_i}$ and $\sum_{i=1}^{\infty} w_i$ converges. It is then a routine exercise to show that

$$\mu = \sum_{i=1}^{\infty} \mu(X) w_i \delta_{x_i} \quad \text{and} \quad \sum_{i=1}^{\infty} \mu(X) w_i \text{ converges.}$$

Finally, suppose that $\mu \in \mathcal{M}(X)$ is arbitrary. Let $\mu = \mu^+ - \mu^-$ be a Jordan decomposition of μ , let $(w_n), (u_n) \in \mathbb{R}$ and let $(x_n), (y_n) \in C_\mu$ be sequences such that $\mu^+ = \sum_{i=1}^{\infty} w_i \delta_{x_i}$ and $\mu^- = \sum_{i=1}^{\infty} u_i \delta_{y_i}$ and both $\sum_{i=1}^{\infty} w_i$ and $\sum_{i=1}^{\infty} u_i$ converge. Subsequently, let $(v_n) \in \mathbb{R}$ and $(z_i) \in C_\mu$ respectively be the sequences $w_1, u_1, w_2, u_2, \dots$

and $x_1, y_1, x_2, y_2, \dots$. It is then a routine exercise to show that

$$\mu = \sum_{i=1}^{\infty} v_i \delta_{z_i} \quad \text{and} \quad \sum_{i=1}^{\infty} v_i \text{ converges.} \quad \square$$

It is worthy of note that in the above proof, for the case that $\mu \in \mathcal{M}^1(X)$, we cannot exclude the possibility that some $w_i = 0$. For example, consider the measure

$$\mu = \sum_{n=1}^{\infty} \frac{1}{2^n} \delta_{1/n} \in \mathcal{M}^1([0, 1]) .$$

Then as every open neighbourhood of 0 contains some point $1/m$, it follows that $0 \in C_\mu$. Hence $x_n = 0$ for some n and $w_n = \mu(\{0\}) = 0$.

Lemma 2.3.23. *Let (X, d) be a totally bounded metric space and let $\varepsilon > 0$. Then there exists a partition $A_1, \dots, A_n \in \mathcal{B}(X)$ of X such that $D(A_i) < \varepsilon$ for each i .*

Proof. As X is totally bounded, there exists $x_1, \dots, x_n \in X$ such that the collection of spheres centred at the x_i of radius $\varepsilon/2$ cover X . Denote each such sphere centred at x_i by U_i , define $A_1 = U_1 \in \mathcal{B}(X)$, and for each $i = 2, \dots, n$, define

$$A_i = U_i \cap \left(X \setminus \bigcup_{j=1}^{i-1} U_j \right) \in \mathcal{B}(X) .$$

Then for each i , $A_i \subseteq U_i$, giving that $D(A_i) \leq D(U_i) \leq \varepsilon$.

It remains to be shown that the A_i partition X . Let $1 \leq i < j \leq n$. Then

$$\begin{aligned} A_i \cap A_j &= U_i \cap \left(X \setminus \bigcup_{k=1}^{i-1} U_k \right) \cap U_j \cap \left(X \setminus \bigcup_{k=1}^{j-1} U_k \right) \\ &= U_j \cap \left(U_i \cap \left(X \setminus \bigcup_{k=1}^{j-1} U_k \right) \right) \\ &= U_j \cap \emptyset \\ &= \emptyset . \end{aligned}$$

Let $x \in X$ and let $i = \min \{i : x \in U_i\}$. Then $x \in U_i \cap \left(X \setminus \bigcup_{j=1}^{i-1} U_j \right) = A_i$, which gives that $X = \bigcup_{i=1}^n A_i$. \square

The proof of the following result is found in Parthasarathy [43, pg. 44].

Theorem 2.3.24. *Let X be a compact metric space. Then the subset of $\mathcal{M}^1(X)$ consisting of the atomic measures is dense in $\mathcal{M}^1(X)$.*

Proof. Let $\mu \in \mathcal{M}^1(X)$ such that μ is not atomic. We want to write μ as the weak-* limit of a sequence of atomic measures.

If the support of μ is countable then by Lemma 2.3.22, there exists sequences $(w_n) \in [0, 1]$ and $(x_n) \in X$ such that $\mu = \sum_{i=1}^{\infty} w_i \delta_{x_i}$ and the series $\sum_{i=1}^{\infty} w_i$ converges. Subsequently, let $(\mu_n) \in \mathcal{M}(X)$ be the sequence defined for $n \in \mathbb{N}$ by

$$\mu_n = \sum_{i=1}^n w_i \delta_{x_i} + \left(\sum_{i=n+1}^{\infty} w_i \right) \delta_{x_{n+1}}.$$

Now

$$\begin{aligned} \|\mu_n - \mu\| &= \left\| \sum_{i=1}^n w_i \delta_{x_i} + \left(\sum_{i=n+1}^{\infty} w_i \right) \delta_{x_{n+1}} - \sum_{i=1}^{\infty} w_i \delta_{x_i} \right\| \\ &= \left\| \left(\sum_{i=n+1}^{\infty} w_i \right) \delta_{x_{n+1}} - \sum_{i=n+1}^{\infty} w_i \delta_{x_i} \right\| \\ &\leq \sum_{i=n+1}^{\infty} w_i \|\delta_{x_{n+1}}\| + \sum_{i=n+1}^{\infty} w_i \|\delta_{x_i}\| \\ &= 2 \sum_{i=n+1}^{\infty} w_i, \end{aligned}$$

which by the convergence of $\sum_{i=1}^{\infty} w_i$ necessarily tends to 0 as $n \rightarrow \infty$. Hence $\mu_n \rightarrow \mu$ and so $\mu_n \rightarrow^* \mu$.

Otherwise, the support of μ is uncountable. Using Lemma 2.3.23, for each $n \in \mathbb{N}$ there exists $m(n) \in \mathbb{N}$ and $A_{n,1}, \dots, A_{n,m(n)} \in \mathcal{B}(X)$ such that the diameter of $A_{n,i}$ is less than $1/n$ for each i , $A_{n,i} \cap A_{n,j} = \emptyset$ for each i and j , and $X = \bigcup_{i=1}^{m(n)} A_{n,i}$. For each n , let $x_{n,1}, \dots, x_{n,m}$ such that $x_{n,i} \in A_{n,i}$ for each $i = 1, \dots, m(n)$, and let

$$\mu_n = \sum_{i=1}^{m(n)} \mu(A_{n,i}) \delta_{x_{n,i}}.$$

Let $f \in C(X)$. We want to show that $\mu_n(f) \rightarrow \mu(f)$. For each n and for each $i = 1, \dots, m(n)$, let

$$\alpha_{n,i} = \inf_{x \in A_{n,i}} f(x) \quad \text{and} \quad \beta_{n,i} = \sup_{x \in A_{n,i}} f(x).$$

Then

$$\begin{aligned}
\left| \int f d\mu_n - \int f d\mu \right| &= \left| \sum_{i=1}^{m(n)} \mu(A_{n,i}) f(x_{n,i}) - \sum_{i=1}^{m(n)} \int_{A_{n,i}} f d\mu \right| \\
&= \left| \sum_{i=1}^{m(n)} \int_{A_{n,i}} f(x_{n,i}) d\mu - \sum_{i=1}^{m(n)} \int_{A_{n,i}} f d\mu \right| \\
&\leq \sum_{i=1}^{m(n)} \int_{A_{n,i}} |f(x_{n,i}) - f| d\mu \\
&\leq \sum_{i=1}^{m(n)} \int_{A_{n,i}} \sup_j (\beta_{n,j} - \alpha_{n,j}) d\mu \\
&= \sup_j (\beta_{n,j} - \alpha_{n,j}).
\end{aligned}$$

As f is uniformly continuous and as the diameter of $A_{n,j}$ converges uniformly in j to 0 as $n \rightarrow \infty$, we have that $\sup_j (\beta_{n,j} - \alpha_{n,j}) \rightarrow 0$ as $n \rightarrow \infty$. Therefore $\mu_n(f) \rightarrow \mu(f)$. \square

Corollary 2.3.25. *Let X be a compact metric space. Then the vector space spanned by $\{\delta_x : x \in X\}$ is dense in $\mathcal{M}(X)$.*

2.4 Summary

This chapter introduced the space $\mathcal{M}(X)$ of finite signed measures on a metric space X , equipped with its weak-* topology. The strong topology on $\mathcal{M}(X)$, which contains the weak-* topology, was briefly mentioned. We made the convention that unless otherwise stated, $\mathcal{M}(X)$ is to be considered as a topological vector space equipped with the weak-* topology.

We investigated convergence and approximation in $\mathcal{M}(X)$ and $\mathcal{M}^1(X)$, and compactness and metrizability in $\mathcal{M}^1(X)$, paying particular attention to the case of X being compact. Many of these properties will be used without explicit reference to the particular result in future work. Our main reference for $\mathcal{M}(X)$ was Choquet [13], and for $\mathcal{M}^1(X)$ was Parthasarathy [43].

Chapter 3

Average distance integrals

Given a compact (connected) metric space (X, d) , Chapter 1 defined the average distances $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$, and noted that these constants may be characterised using (signed) Borel measures of unit mass. This chapter introduces two real-valued functions defined by integrating a metric, namely $d_\mu: X \rightarrow \mathbb{R}$ and $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$, and studies properties of each. In particular, we determine if d_μ and I are continuous when $X = [0, 1]$. Chapter 4 will subsequently use d_μ and I to give measure theoretic characterisations of $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$, and to investigate certain properties of these average distances.

3.1 The function $d_\mu: X \rightarrow \mathbb{R}$

Let (X, d) be a compact metric space. For all $y \in X$, the mapping $X \rightarrow \mathbb{R}$ such that $x \mapsto d(x, y)$ is bounded and continuous, and hence integrable with respect to any finite signed Borel measure. Subsequently, for each $\mu \in \mathcal{M}(X)$, let $d_\mu: X \rightarrow \mathbb{R}$ be the function defined for $y \in X$ by

$$d_\mu(y) = \int d(x, y) d\mu(x).$$

3.1.1 Properties of d_μ

We shall now show that $d_\mu \in C(X)$ for each $\mu \in \mathcal{M}(X)$, and give a sufficient condition for a sequence of such functions to be uniformly convergent in $C(X)$.

Proposition 3.1.1. *Let (X, d) be a compact metric space and let $\mu \in \mathcal{M}(X)$. Then $d_\mu \in C(X)$.*

Proof. As a compact metric space has finite diameter, $D(X) < \infty$. Then for all $y \in X$,

$$\begin{aligned} |d_\mu(y)| &= \left| \int d(x, y) d\mu(y) \right| \\ &\leq \int d(x, y) d|\mu|(y) \\ &\leq \int D(X) d|\mu|(y) \\ &= D(X) \|\mu\| \\ &< \infty. \end{aligned}$$

Therefore, d_μ is bounded.

Let $\varepsilon > 0$, let $\delta = \varepsilon/(1 + \|\mu\|) > 0$ and let $y_1, y_2 \in X$ such that $d(y_1, y_2) < \delta$. Then

$$\begin{aligned} |d_\mu(y_1) - d_\mu(y_2)| &= \left| \int d(x, y_1) d\mu(x) - \int d(x, y_2) d\mu(x) \right| \\ &\leq \int |d(x, y_1) - d(x, y_2)| d|\mu|(x) \\ &\leq \int d(y_1, y_2) d|\mu|(x) \\ &< \int \delta d|\mu|(x) \\ &= \frac{\varepsilon \|\mu\|}{1 + \|\mu\|} \\ &< \varepsilon. \end{aligned}$$

Therefore, d_μ is uniformly continuous, and it follows that $d_\mu \in C(X)$. □

The following result generalises a lemma of Wolf [56, pg. 396].

Proposition 3.1.2. *Let (X, d) be a compact metric space and let $(\mu_\alpha) \in \mathcal{M}(X)$ be a convergent net with limit $\mu \in \mathcal{M}(X)$. If $(\|\mu_\alpha\|)$ is bounded then $(d_{\mu_\alpha}) \in C(X)$ converges uniformly to d_μ .*

Proof. Suppose that $(\|\mu_\alpha\|)$ is bounded. Then there exists $r > 0$ such that $\|\mu_\alpha\| \leq r$ for all α , and $\|\mu\| \leq r$. Let $\varepsilon > 0$ and let $\delta = \varepsilon/3r$. As X is compact, there exists a finite subset $\{x_1, \dots, x_m\}$ of X such that the collection of open spheres centred at the x_i of radius δ covers X . For all $y \in X$, the mapping $X \rightarrow \mathbb{R}$ such that $x \mapsto d(x, y)$ is bounded and continuous, and as $\mu_\alpha \rightarrow^* \mu$ it follows that

$$d_{\mu_\alpha}(y) = \int d(x, y) d\mu_\alpha(x) \rightarrow \int d(x, y) d\mu(x) = d_\mu(y).$$

That is, (d_{μ_α}) converges pointwise to d_μ . In particular, there exists an α_0 such that for each $i = 1, \dots, m$, if $\alpha_0 < \alpha$ then $|d_{\mu_\alpha}(x_i) - d_\mu(x_i)| < \varepsilon/3$.

Let $y \in X$. Then $d(y, x_i) < \delta$ for some $i = 1, \dots, m$ and $\alpha_0 < \alpha$ implies that

$$\begin{aligned} |(d_{\mu_\alpha} - d_\mu)(y)| &\leq |d_{\mu_\alpha}(y) - d_{\mu_\alpha}(x_i)| + |d_{\mu_\alpha}(x_i) - d_\mu(x_i)| + |d_\mu(x_i) - d_\mu(y)| \\ &\leq \int d(y, x_i) d|\mu_\alpha|(x) + \frac{\varepsilon}{3} + \int d(y, x_i) d|\mu|(x) \\ &= d(y, x_i) \|\mu_\alpha\| + \frac{\varepsilon}{3} + d(y, x_i) \|\mu\| \\ &< \delta r + \frac{\varepsilon}{3} + \delta r \\ &= \varepsilon. \end{aligned}$$

It follows from the compactness of X that if $\alpha_0 < \alpha$ then

$$\|d_{\mu_\alpha} - d_\mu\| = \sup_{y \in X} |(d_{\mu_\alpha} - d_\mu)(y)| < \varepsilon.$$

Therefore (d_{μ_α}) converges to d_μ in $C(X)$. □

The original result of Wolf is:

Corollary 3.1.3. *Let (X, d) be a compact metric space and let $(\mu_n) \in \mathcal{M}^1(X)$ be a convergent sequence with limit $\mu \in \mathcal{M}^1(X)$. Then $(d_{\mu_n}) \in C(X)$ converges uniformly to d_μ .*

3.2 The functional $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$

Let (X, d) be a compact metric space. We denote the product space $\mathcal{M}(X) \times \mathcal{M}(X)$ by $\mathcal{M}(X)^2$, and we denote the subspace of $\mathcal{M}(X)^2$ which consists precisely of all pairs of probability measures by $\mathcal{M}^1(X)^2$. We also use the symbol \rightarrow^* to denote coordinate-wise weak- $*$ convergence in $\mathcal{M}(X)^2$ and $\mathcal{M}^1(X)^2$, which is of course convergence in the respective product topologies.

Now by Proposition 3.1.1, d_μ is bounded and continuous and hence integrable when $\mu \in \mathcal{M}(X)$. Subsequently, let $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ be the functional defined for $(\mu, \nu) \in \mathcal{M}(X)^2$ by

$$I(\mu, \nu) = \int d_\mu d\nu = \iint d(x, y) d\mu(x) d\nu(y).$$

The abbreviation $I(\mu)$ will be used for $I(\mu, \mu)$ where convenient. In this case, I may be considered as a functional $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ defined for $\mu \in \mathcal{M}(X)$ by

$$I(\mu) = \int d_\mu d\mu = \iint d(x, y) d\mu(x) d\mu(y).$$

As we will consider the restrictions of $I(\cdot, \cdot)$ and $I(\cdot)$ to $\mathcal{M}^1(X)^2$ and $\mathcal{M}^1(X)$, the symbol I may be used to represent any one of four functionals, the exact meaning of which shall be made clear by context.

Let $\mu, \nu \in \mathcal{M}(X)$. Then by Fubini's Theorem, $I(\mu, \nu) = I(\nu, \mu)$, and using linearity properties of the integral we have the identities

$$\begin{aligned} I(\mu + \nu) &= I(\mu) + I(\nu) + 2I(\mu, \nu), \\ I(\mu - \nu) &= I(\mu) + I(\nu) - 2I(\mu, \nu). \end{aligned}$$

These identities will be used in future work without reference.

3.2.1 Continuity properties of $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$

We shall see by counterexample in Section 3.2.2 that if X is a compact metric space then the functional $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ need not be continuous, in which case it is

discontinuous at each point. For now, we will show that I is continuous on various important subsets of $\mathcal{M}(X)^2$.

The following result is a generalisation of a lemma due to Wolf [56, pg. 396].

Theorem 3.2.1. *Let (X, d) be a compact metric space and let B be a closed ball in $\mathcal{M}(X)$ of finite radius. Then the restriction of $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ to B is bounded and continuous with respect to the weak- $*$ topology.*

Proof. Let $r \geq 0$ and let $B = \{\mu \in \mathcal{M}(X) : \|\mu\| \leq r\}$. It will be sufficient to show that I is bounded and continuous when restricted to B . For all $\mu \in B$,

$$\begin{aligned} |I(\mu)| &= \left| \iint d(x, y) d\mu(x) d\mu(y) \right| \\ &\leq \iint D(X) d|\mu|(x) d|\mu|(y) \\ &= D(X) \|\mu\|^2 \\ &\leq D(X) r^2 \\ &< \infty. \end{aligned}$$

Therefore, I is bounded on B .

Let $(\mu_\alpha) \in B$ be a convergent net with limit $\mu \in B$. By Proposition 3.1.1 we have that $d_\mu \in C(X)$, giving $\mu_\alpha(d_\mu) \rightarrow \mu(d_\mu)$. Also, by Proposition 3.1.2, $\|d_{\mu_\alpha} - d_\mu\| \rightarrow 0$, and we then have

$$\begin{aligned} |I(\mu_\alpha) - I(\mu)| &\leq \left| \int d_{\mu_\alpha} d\mu_\alpha - \int d_\mu d\mu_\alpha \right| + \left| \int d_\mu d\mu_\alpha - \int d_\mu d\mu \right| \\ &\leq \int \|d_{\mu_\alpha} - d_\mu\| d|\mu_\alpha| + \left| \int d_\mu d\mu_\alpha - \int d_\mu d\mu \right| \\ &= \|d_{\mu_\alpha} - d_\mu\| \cdot \|\mu_\alpha\| + \left| \int d_\mu d\mu_\alpha - \int d_\mu d\mu \right| \\ &\leq \|d_{\mu_\alpha} - d_\mu\| \cdot r + \left| \int d_\mu d\mu_\alpha - \int d_\mu d\mu \right| \\ &\rightarrow 0 \cdot r + 0 \\ &= 0. \end{aligned}$$

Therefore, I is continuous on B . □

The identity $I(\mu, \nu) = \frac{1}{2}(I(\mu + \nu) - I(\mu) - I(\nu))$ in $\mathcal{M}(X)$ gives:

Corollary 3.2.2. *Let X be a compact metric space and let B be a closed ball in $\mathcal{M}(X)$ of finite radius. Then the restriction of $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ to $B \times B$ is bounded and continuous with respect to the weak-* topology on $B \times B$. Further, $I: \mathcal{M}^1(X)^2 \rightarrow \mathbb{R}$ is bounded and continuous.*

The original result of Wolf is:

Corollary 3.2.3. *Let X be a compact metric space. Then $I: \mathcal{M}^1(X) \rightarrow \mathbb{R}$ is bounded and continuous.*

We now turn to a discussion of the continuity I on its entire domain.

Proposition 3.2.4. *Let X be a compact metric space. Then $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ is separately continuous.*

Proof. Let $(\nu_\alpha) \in \mathcal{M}(X)$ be a convergent net and let $\mu, \nu \in \mathcal{M}(X)$ such that $\nu_\alpha \rightarrow^* \nu$. As I is symmetric, it is sufficient to show that $I(\mu, \nu_\alpha) \rightarrow I(\mu, \nu)$. Recalling from Proposition 3.1.1 that $d_\mu \in C(X)$, this is immediate from the usual characterisation of weak-* convergence in $\mathcal{M}(X)$, and noting that $I(\mu, \nu_\alpha) = \nu_\alpha(\mu)$ and $I(\mu, \nu) = \nu(\mu)$. □

Proposition 3.2.5. *Let X be a compact metric space. Then $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ is continuous at each point if and only if it is continuous at one point.*

Proof. The necessary condition is obvious. To prove the sufficient condition, suppose that I is continuous at some point $(\mu_0, \nu_0) \in \mathcal{M}(X)^2$. Let $(\mu_\alpha, \nu_\alpha) \in \mathcal{M}(X)^2$ be a convergent net with limit $(\mu, \nu) \in \mathcal{M}(X)^2$. Then as addition of signed measures is continuous, $(\mu_\alpha - \mu + \mu_0, \nu_\alpha - \nu + \nu_0) \rightarrow^* (\mu_0, \nu_0)$, which gives that

$$I(\mu_\alpha - \mu + \mu_0, \nu_\alpha - \nu + \nu_0) \rightarrow I(\mu_0, \nu_0).$$

Now,

$$\begin{aligned}
& I(\mu_\alpha - \mu + \mu_0, \nu_\alpha - \nu + \nu_0) \\
&= I(\mu_\alpha - \mu, \nu_\alpha - \nu) + I(\mu_\alpha - \mu, \nu_0) + I(\mu_0, \nu_\alpha - \nu) + I(\mu_0, \nu_0) \\
&= I(\mu_\alpha, \nu_\alpha) - I(\mu_\alpha, \nu) - I(\mu, \nu_\alpha) + I(\mu, \nu) \\
&\quad + I(\mu_\alpha - \mu, \nu_0) + I(\mu_0, \nu_\alpha - \nu) + I(\mu_0, \nu_0).
\end{aligned}$$

Using Proposition 3.2.4, we have that

$$\begin{aligned}
I(\mu_\alpha, \nu) &\rightarrow I(\mu, \nu), \\
I(\mu, \nu_\alpha) &\rightarrow I(\mu, \nu), \\
I(\mu_\alpha - \mu, \nu_0) &\rightarrow I(0, \nu_0) = 0, \\
I(\mu_0, \nu_\alpha - \nu_0) &\rightarrow I(\mu_0, 0) = 0.
\end{aligned}$$

It follows that $I(\mu_\alpha, \nu_\alpha) \rightarrow I(\mu, \nu)$, and I is continuous at each point. \square

Corollary 3.2.6. *Let X be a compact metric space. Then $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ is discontinuous at each point if and only if it is discontinuous at one point.*

Proposition 3.2.7. *Let X be a compact metric space. Then $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ is continuous at each point if and only if it is continuous at one point.*

Proof. The necessary condition is obvious. To prove the sufficient condition, suppose that I is continuous at some point $\mu_0 \in \mathcal{M}(X)$. Let $(\mu_\alpha) \in \mathcal{M}(X)$ be a convergent net with limit $\mu \in \mathcal{M}(X)$. Then as addition of signed measures is continuous, $\mu_\alpha - \mu + \mu_0 \rightarrow^* \mu_0$, which gives that $I(\mu_\alpha - \mu + \mu_0) \rightarrow I(\mu_0)$. Now, \bullet

$$\begin{aligned}
& I(\mu_\alpha - \mu + \mu_0) \\
&= I(\mu_\alpha - \mu) + I(\mu_0) + 2I(\mu_\alpha - \mu, \mu_0) \\
&= I(\mu_\alpha) + I(\mu) + I(\mu_0) - 2I(\mu_\alpha, \mu) + 2I(\mu_\alpha - \mu, \mu_0)
\end{aligned}$$

Using Proposition 3.2.4, we have that

$$\begin{aligned}
I(\mu_\alpha, \mu) &\rightarrow I(\mu) \\
I(\mu_\alpha - \mu, \mu_0) &\rightarrow I(0, \mu_0) = 0.
\end{aligned}$$

It follows that $I(\mu_\alpha) \rightarrow I(\mu)$, and I is continuous at each point. \square

Corollary 3.2.8. *Let X be a compact metric space. Then $I: \mathcal{M}^1(X)^2 \rightarrow \mathbb{R}$ is discontinuous at each point if and only if it is discontinuous at one point.*

We now work towards a sufficient condition to ensure that $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ and $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ are discontinuous. For each compact metric space X , and for each $f_1, \dots, f_n \in C(X)$, let

$$V(f_1, \dots, f_n) = \left\{ \mu \in \mathcal{M}(X) : \int f_i d\mu = 0 \text{ for } i = 1, \dots, n \right\}.$$

The following three lemmas are derived from exercises in Bourbaki [10, pg.104].

The proofs are due to Peter Nickolas.

Lemma 3.2.9. *Let V be a linear space and let $v_1, \dots, v_n \in V$ be independent. Then there exist linear transformations $\phi_i: V \rightarrow \mathbb{R}$ for $i = 1, \dots, n$ such that the matrix $(\phi_i(v_j))$ is non-singular.*

Proof. Let $U = \langle v_1, \dots, v_n \rangle$ be the subspace of V spanned by the v_i and let W be a subspace of V such that $V = U \oplus W$. Then U is a finite dimensional subspace of V , and each $v \in V$ has a unique representation $v = \sum_{i=1}^n \alpha_i v_i + w$ where $\sum_{i=1}^n \alpha_i v_i \in U$ and $w \in W$. For each $i = 1, \dots, n$, let $\phi_i: V \rightarrow \mathbb{R}$ be the function defined by $\phi_i(v) = \alpha_i$. It is clear that each ϕ_i is a linear transformation. Now, for each i and for each $v \in V$, $\phi_i(v)$ is the i -th component of the projection of v onto U . It follows that the j -th column of $(\phi_i(v_j))$ is simply the coordinate vector of v_j in U , giving that $(\phi_i(v_j))$ is non-singular. \square

Lemma 3.2.10. *Let X be a compact metric space and let $f_1, \dots, f_n \in C(X)$. Then $V(f_1, \dots, f_n)$ is a subspace of $\mathcal{M}(X)$ and $\text{codim}(V(f_1, \dots, f_n)) \leq n$, with equality occurring if and only if the f_i are independent.*

Proof. Let $V = V(f_1, \dots, f_n)$. Using linearity properties of the integral, V is a subspace of $\mathcal{M}(X)$. Hence, let W be a subspace of V such that $\mathcal{M}(X) = V \oplus W$.

Suppose that the f_i are independent. Then by Lemma 3.2.9 and Theorem 2.3.3 there exist linear functionals $\mu_1, \dots, \mu_n \in \mathcal{M}(X)$ such that $(\mu_i(f_j))$ is non-singular.

As $(\mu_i(f_j))$ is non-singular, it follows that the μ_i are independent and each $\mu_i \notin V$, giving $\mu_i \in W$. Now, let $\mu \in \mathcal{M}(X)$ and consider the system of linear equations in $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ given by

$$\mu(f_i) + \sum_{j=1}^n \mu_j(f_i) \alpha_j = 0, \quad i = 1, \dots, n.$$

This system has coefficient matrix $(\mu_j(f_i))$, which is necessarily non-singular, and hence has a unique solution. Using linearity properties of the integral, let $\alpha_1, \dots, \alpha_n$ be the unique real numbers such that

$$\int f_i d\left(\mu + \sum_{j=1}^n \alpha_j \mu_j\right) = 0, \quad i = 1, \dots, n,$$

let $\nu = \mu + \sum_{i=1}^n \alpha_i \mu_i \in V$ and let $\omega = -\sum_{i=1}^n \alpha_i \mu_i \in W$. Then $\mu = \nu + \omega$ is the unique decomposition of μ where $\nu \in V$ and $\omega \in W$. Now, if $\mu \in W$ then $\nu = 0$ and $\omega = \mu$. Hence, by varying μ , the mapping $\mathcal{M}(X) \rightarrow W$ such that $\mu \mapsto \omega$ is onto, and it follows that W is spanned by the independent μ_i . Therefore, $\dim(W) = n$ and $\text{codim}(V) = n$.

Otherwise, the f_i are dependent. Select g_1, \dots, g_m from the f_i such that the g_i are independent and $m < n$ is maximal, and let $V' = V(g_1, \dots, g_m)$. Using linearity properties of the integral, $V' = V$ and by the above, $\text{codim}(V) = m < n$. \square

Lemma 3.2.11. *Let X be a compact metric space such that $\{d_\mu: \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$ and let $f_1, \dots, f_n \in C(X)$ be independent. Then for some $\mu \in \mathcal{M}(X)$, f_1, \dots, f_n, d_μ are independent.*

Proof. Suppose that for all $\mu \in \mathcal{M}(X)$, f_1, \dots, f_n, d_μ are dependent. Then as the f_i are independent, each d_μ can be expressed as a linear combination of the f_i . As $\{d_\mu: \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$, it follows that the finite-dimensional subspace $\langle f_1, \dots, f_n \rangle \subseteq C(X)$ is dense in the infinite dimensional space $C(X)$, which is a contradiction. Therefore, f_1, \dots, f_n, d_μ are independent for some $\mu \in \mathcal{M}(X)$. \square

The following result is due to Peter Nickolas.

Theorem 3.2.12. *Let X be a compact metric space such that $\{d_\mu : \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$. Then there exists a convergent net $(\mu_\alpha, \nu_\alpha) \in \mathcal{M}(X)^2$ with limit $(\mu, 0) \in \mathcal{M}(X)^2$ such that $\|\mu_\alpha\| = 1$ for all α and $I(\mu_\alpha, \nu_\alpha) \not\rightarrow I(\mu, 0) = 0$.*

Proof. Let A be the directed set of all finite subsets of $C(X)$ other than $\{0\}$ equipped with set inclusion, and let $\alpha = \{f_1, \dots, f_n\} \in A$. For convenience, we will denote by $V(\alpha)$ the set $V(f_1, \dots, f_n)$.

By Proposition 3.2.11, there exists $\mu_\alpha \in \mathcal{M}(X)$ such that d_{μ_α} is not in the linear span of α . Now, for all $c \in \mathbb{R}$ such that $c \neq 0$, $d_{c\mu_\alpha} = cd_{\mu_\alpha}$, and it follows that $d_{c\mu_\alpha}$ is not in the linear span of α . We may then assume without loss of generality that μ_α has been chosen such that $\|\mu_\alpha\| = 1$. Now, let $B = \{\mu \in \mathcal{M}(X) : \|\mu\| \leq 1\}$ be the closed unit ball in $\mathcal{M}(X)$. Then by Proposition 2.3.20, B is compact and as $(\mu_\alpha) \in B$, there exists a subnet $(\mu_{\alpha(\beta)})$ of (μ_α) and $\mu \in B$ such that $\mu_{\alpha(\beta)} \rightarrow^* \mu$.

Let $\alpha' \subseteq \alpha$ be a maximal independent set. Then Proposition 3.2.10 gives that $\text{codim}(V(\alpha)) = |\alpha'|$ and $\text{codim}(V(\alpha \cup \{d_{\mu_\alpha}\})) = |\alpha'| + 1$, and so there must exist some $\nu_\alpha \in V(\alpha)$ such that $\nu_\alpha \notin V(\alpha \cup \{d_{\mu_\alpha}\})$. We have that for each $f \in V(\alpha)$, $\nu_\alpha(f) = 0$, and $I(\mu_\alpha, \nu_\alpha) = \nu_\alpha(d_{\mu_\alpha}) \neq 0$. As $V(\alpha)$ and $V(\alpha \cup \{d_{\mu_\alpha}\})$ are vector spaces, we may assume without loss of generality that ν_α has been chosen such that $I(\mu_\alpha, \nu_\alpha) = |\alpha|$. Then for all $f \in C(X)$ and for all $\alpha \in A$,

$$\{f\} \subseteq \alpha \Rightarrow f \in \alpha \Rightarrow \int f d\nu_\alpha = 0.$$

Therefore, $\nu_\alpha \rightarrow^* 0$ and $\nu_{\alpha(\beta)} \rightarrow^* 0$.

We then have that in $\mathcal{M}(X)^2$, $(\mu_{\alpha(\beta)}, \nu_{\alpha(\beta)}) \rightarrow^* (\mu, 0)$, but

$$I(\mu_{\alpha(\beta)}, \nu_{\alpha(\beta)}) = |\alpha(\beta)| \not\rightarrow 0 = I(\mu, 0).$$

The result follows by renaming the net $(\mu_{\alpha(\beta)}, \nu_{\alpha(\beta)})$ to (μ_α, ν_α) . □

Corollary 3.2.13. *Let X be a compact metric space such that $\{d_\mu : \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$. Then $I : \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ is discontinuous at each point.*

Proof. The result follows from Proposition 3.2.5 and Proposition and 3.2.12. □

Corollary 3.2.14. *Let X be a compact metric space such that $\{d_\mu: \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$. Then $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ is discontinuous at each point.*

Proof. The result follows from Corollary 3.2.13, the identity

$$I(\mu, \nu) = \frac{I(\mu + \nu) - I(\mu) - I(\nu)}{2}$$

for $\mu, \nu \in \mathcal{M}(X)$, and the continuity of addition in $\mathcal{M}(X)$. □

We conclude this section by noting the following result of Nickolas and Wolf [40]. The strict quasihypermetric property is discussed in Chapter 5.

Theorem 3.2.15. *Let X be a strictly quasihypermetric compact metric space. Then each of $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ and $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ are continuous if and only if X is finite.*

3.2.2 Discontinuity of $I: \mathcal{M}([0, 1])^2 \rightarrow \mathbb{R}$

We now show that $I: \mathcal{M}([0, 1])^2 \rightarrow \mathbb{R}$ is discontinuous at each point. As a result of Corollary 3.2.13, it will be sufficient to show that $\{d_\mu: \mu \in C([0, 1])\}$ is dense in $C([0, 1])$. Subsequently, we investigate the mapping $\mathcal{M}([0, 1]) \rightarrow C([0, 1])$ such that $\mu \mapsto d_\mu$.

Lemma 3.2.16. *The collection of all piecewise linear continuous functions on $[0, 1]$ is dense in $C([0, 1])$.*

Proof. Let $f \in C([0, 1])$ and let $\varepsilon > 0$. By the compactness of $[0, 1]$, f is uniformly continuous, and there exists $\delta > 0$ such that for all $x_1, x_2 \in [0, 1]$, $|x_1 - x_2| < \delta$ implies that $|f(x_1) - f(x_2)| < \varepsilon/2$.

Let $0 = a_0 < a_1 < \dots < a_{n-1} < a_n = 1$ be a subdivision of $[0, 1]$ into intervals of equal length such that $|a_0 - a_1| < \delta$ and let $g: [0, 1] \rightarrow \mathbb{R}$ be the function defined for $x \in [a_i, a_{i+1}]$ by

$$g(x) = f(a_i) + \frac{f(a_{i+1}) - f(a_i)}{a_{i+1} - a_i}(x - a_i).$$

Then g is a piecewise linear continuous function on $[0, 1]$ such that $g(a_i) = f(a_i)$ for each i . Let $x \in [0, 1]$. Then $x \in [a_i, a_{i+1}]$ for some $i = 0, \dots, n-1$ and

$$\begin{aligned} |g(a_i) - g(x)| &= \left| f(a_i) - \left(f(a_i) + \frac{f(a_{i+1}) - f(a_i)}{a_{i+1} - a_i} (x - a_i) \right) \right| \\ &= |f(a_{i+1}) - f(a_i)| \cdot \left| \frac{x - a_i}{a_{i+1} - a_i} \right| \\ &\leq |f(a_{i+1}) - f(a_i)| \\ &< \varepsilon/2. \end{aligned}$$

It follows that

$$|(f - g)(x)| \leq |f(x) - f(a_i)| + |g(a_i) - g(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

We have that $|(f - g)(x)| < \varepsilon$ for all $x \in [0, 1]$, and as $[0, 1]$ is compact, it follows that

$$\|f - g\| = \sup_{x \in [0, 1]} |(f - g)(x)| < \varepsilon.$$

Hence, there exists a sequence (g_n) of piecewise linear continuous functions on $[0, 1]$ such that $\|f - g_n\| \rightarrow 0$. □

The following result is derived from an exercise in Bourbaki [10, pg.103]. The proof is due to Peter Nickolas.

Theorem 3.2.17. *Let $\phi: \mathcal{M}([0, 1]) \rightarrow C([0, 1])$ be the mapping defined for each $\mu \in \mathcal{M}([0, 1])$ by $\phi(\mu) = d_\mu$. Then*

1. $\phi(\mathcal{M}([0, 1]))$ is dense in $C([0, 1])$.
2. ϕ is injective.
3. ϕ is continuous with respect to the norm topology on $\mathcal{M}([0, 1])$.
4. ϕ is discontinuous at each point with respect to the weak-* topology on $\mathcal{M}([0, 1])$.
5. ϕ^{-1} is discontinuous at each point with respect to the weak-* topology on $\mathcal{M}([0, 1])$.

Proof. 1. We firstly show that the image of $\mathcal{M}([0, 1])$ under ϕ is dense in $C([0, 1])$.

Let $\mu = \delta_0 + \delta_1 \in \mathcal{M}([0, 1])$. Then for all $y \in [0, 1]$,

$$(\phi(\mu))(y) = \int |x - y| d\mu(x) = |0 - y| + |1 - y| = y + 1 - y = 1.$$

Therefore, $\phi(\mathcal{M}([0, 1]))$ contains the mapping $[0, 1] \rightarrow \mathbb{R}$ such that $x \mapsto 1$.

Let $a \in [0, 1]$ and let $\mu = \delta_a \in \mathcal{M}([0, 1])$. Then for all $y \in [0, 1]$,

$$(\phi(\mu))(y) = \int |x - y| d\mu(x) = |a - y| = |y - a|.$$

Therefore, $\phi(\mathcal{M}([0, 1]))$ contains the mapping $[0, 1] \rightarrow \mathbb{R}$ such that $x \mapsto |x - a|$.

For each $0 \leq a < b \leq 1$, let $\psi_{a,b}: [0, 1] \rightarrow \mathbb{R}$ be the function defined for $x \in [0, 1]$ by

$$\psi_{a,b}(x) = \frac{1}{2} + \frac{1}{2(b-a)} (|x - a| - |x - b|).$$

Using properties of the integral, it is a routine exercise to show that $\phi(\mathcal{M}([0, 1]))$ is a vector subspace of $C([0, 1])$. Hence $\psi_{a,b} \in \phi(\mathcal{M}([0, 1]))$ for all $0 \leq a < b \leq 1$. To show that $\phi(\mathcal{M}([0, 1]))$ is dense in $C([0, 1])$, it will be sufficient to show that any piecewise linear continuous function on $[0, 1]$ is a linear combination of such functions $\psi_{a,b}$. Let $f: [0, 1] \rightarrow \mathbb{R}$ be a piecewise linear continuous function on $[0, 1]$. Then there exists $0 = a_0 < a_1 < \dots < a_{n-1} < a_n = 1 \in [0, 1]$ such that for all $x \in [a_i, a_{i+1}]$

$$f(x) = f(a_i) + \frac{f(a_{i+1}) - f(a_i)}{a_{i+1} - a_i} (x - a_i).$$

We now claim that

$$f = f(a_0) + \sum_{i=0}^{n-1} (f(a_{i+1}) - f(a_i)) \psi_{a_i, a_{i+1}}.$$

Let $x \in [0, 1]$. Then $x \in [a_i, a_{i+1}]$ for some $i = 0, \dots, n-1$. It will be sufficient to show that

$$f(a_0) + \sum_{j=0}^{n-1} (f(a_{j+1}) - f(a_j)) \psi_{a_j, a_{j+1}}(x) = f(a_i) + \frac{f(a_{i+1}) - f(a_i)}{a_{i+1} - a_i} (x - a_i).$$

Now, for all $j = 0, \dots, i-1$, a_{j+1} lies on the line segment with endpoints x and a_j . Therefore $|x - a_j| = |x - a_{j+1}| + |a_{j+1} - a_j|$ and

$$\begin{aligned}\psi_{a_j, a_{j+1}}(x) &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (|x - a_j| - |x - a_{j+1}|) \\ &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (|x - a_{j+1}| + |a_{j+1} - a_j| - |x - a_{j+1}|) \\ &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (a_{j+1} - a_j) \\ &= 1.\end{aligned}$$

Similarly, for all $j = i+1, \dots, n$, a_j lies on the line segment with endpoints x and a_{j+1} . Therefore $|x - a_{j+1}| = |x - a_j| + |a_j - a_{j+1}|$ and

$$\begin{aligned}\psi_{a_j, a_{j+1}}(x) &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (|x - a_j| - |x - a_{j+1}|) \\ &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (|x - a_j| - (|x - a_j| + |a_j - a_{j+1}|)) \\ &= \frac{1}{2} + \frac{1}{2(a_{j+1} - a_j)} (a_j - a_{j+1}) \\ &= 0.\end{aligned}$$

We evaluate $\sum_{j=0}^{i-1} (f(a_{j+1}) - f(a_j)) = f(a_i) - f(a_0)$ as a telescoping sum. Therefore,

$$\begin{aligned}& f(a_0) + \sum_{j=0}^{n-1} (f(a_{j+1}) - f(a_j)) \psi_{a_j, a_{j+1}}(x) \\ &= f(a_0) + \sum_{j=0}^{i-1} (f(a_{j+1}) - f(a_j)) + (f(a_{i+1}) - f(a_i)) \psi_{a_i, a_{i+1}}(x) \\ &= f(a_i) + (f(a_{i+1}) - f(a_i)) \left(\frac{1}{2} + \frac{1}{2(a_{i+1} - a_i)} (|x - a_i| - |x - a_{i+1}|) \right) \\ &= f(a_i) + (f(a_{i+1}) - f(a_i)) \cdot \frac{a_{i+1} - a_i + (x - a_i) - (a_{i+1} - x)}{2(a_{i+1} - a_i)} \\ &= f(a_i) + \frac{f(a_{i+1}) - f(a_i)}{a_{i+1} - a_i} (x - a_i).\end{aligned}$$

It follows that $\phi(\mathcal{M}([0, 1]))$ is dense in $C([0, 1])$.

2. We now show that ϕ is injective. Let $\mu, \nu \in \mathcal{M}([0, 1])$. Then $\phi(\mu) = \phi(\nu)$ if and only if $d_{\mu-\nu} = 0$. Suppose that $d_\mu = 0$. It will then be sufficient to show that

$\mu = 0$, which we shall establish in three steps. Firstly, note that for all $y \in [0, 1]$,

$$\int |x - y| d\mu(x) = 0,$$

Then for all $c \in \mathbb{R}$,

$$\begin{aligned} \int c d\mu(x) &= c \int 1 d\mu(x) \\ &= c \int (|x - 0| + |x - 1|) d\mu(x) \\ &= c \int |x - 0| d\mu(x) + c \int |x - 1| d\mu(x) \\ &= 0. \end{aligned}$$

Let $0 \leq a < b \leq 1$. By linearity of the integral, it follows that $\mu(\phi_{a,b}) = 0$. Secondly, let $f: [0, 1] \rightarrow \mathbb{R}$ be a piecewise linear continuous function on $[0, 1]$. Then there exists $0 = a_0 < a_1 < \dots < a_{n-1} < a_n = 1 \in [0, 1]$ such that

$$f = f(a_0) + \sum_{i=0}^{n-1} (f(a_{i+1}) - f(a_i)) \phi_{a_i, a_{i+1}},$$

and again by linearity of the integral, $\mu(f) = 0$. Finally, let $f \in C([0, 1])$ and let $(f_n) \in C([0, 1])$ be a sequence of piecewise linear functions such that $f_n \rightarrow f$. Then $\mu(f_n) \rightarrow \mu(f)$, giving that $\mu(f) = 0$. As $\mu(f) = 0$ for all $f \in C([0, 1])$, we then have by Corollary 2.3.8 that $\mu = 0$. Hence, ϕ is injective.

3. We now show that ϕ is continuous with respect to the norm topology on $\mathcal{M}(X)$. Let $(\mu_\alpha) \in \mathcal{M}([0, 1])$ be a strongly convergent net with limit $\mu \in \mathcal{M}([0, 1])$. Then

$$\begin{aligned} \|d_{\mu_\alpha} - d_\mu\| &= \sup_{y \in [0, 1]} \left| \int |x - y| d\mu_\alpha(x) - \int |x - y| d\mu(x) \right| \\ &= \sup_{y \in [0, 1]} \left| \int |x - y| d(\mu_\alpha - \mu)(x) \right| \\ &\leq \sup_{y \in [0, 1]} \int |x - y| d|\mu_\alpha - \mu|(x) \\ &\leq \int 1 d|\mu_\alpha - \mu|(x) \\ &= \|\mu_\alpha - \mu\| \\ &\rightarrow 0. \end{aligned}$$

Therefore $d_{\mu_\alpha} \rightarrow d_\mu$ in $C([0, 1])$, and ϕ is continuous with respect to the norm topology on $\mathcal{M}([0, 1])$.

4. Note that Part 1 may be stated as $\{d_\mu: \mu \in \mathcal{M}([0, 1])\}$ is dense in $C([0, 1])$. By Proposition 3.2.12, there exists a convergent net $(\mu_\alpha, \nu_\alpha) \in \mathcal{M}([0, 1])^2$ with limit $(\mu, 0) \in \mathcal{M}([0, 1])^2$ such that $\|\mu_\alpha\| = 1$ for all α and $I(\mu_\alpha, \nu_\alpha) \not\rightarrow I(\mu, 0) = 0$. Now, suppose that ϕ is continuous at $0 \in \mathcal{M}([0, 1])$. Then as $\nu_\alpha \rightarrow^* 0$, it follows that $\|d_{\nu_\alpha} - d_0\| = \|d_{\nu_\alpha}\| \rightarrow 0$ and

$$|I(\mu_\alpha, \nu_\alpha)| = \left| \int d_{\nu_\alpha} d\mu_\alpha \right| \leq \|d_{\nu_\alpha}\| \cdot \|\mu_\alpha\| = \|d_{\nu_\alpha}\| \rightarrow 0.$$

It follows that $I(\mu_\alpha, \nu_\alpha) \rightarrow 0$, which is a contradiction. Therefore ϕ is not continuous at 0. By linearity properties of the integral, ϕ is discontinuous with respect to the weak-* topology on $\mathcal{M}([0, 1])$ at each point if and only if it is discontinuous at one point. The result follows.

5. It remains to be shown that ϕ^{-1} is not continuous with respect to the weak-* topology on $\mathcal{M}(X)$. For each $n \in \mathbb{N}$, let

$$\mu_n = (n+1)\delta_0 - n\delta_{\frac{1}{n^2}} \in \mathcal{M}([0, 1]).$$

Then

$$\begin{aligned} \|\phi(\mu_n) - \phi(\delta_0)\| &= \sup_{y \in [0, 1]} \left| \int |x - y| d\mu_n(x) - \int |x - y| d\delta_0(x) \right| \\ &= \sup_{y \in [0, 1]} \left| \int |x - y| d(n\delta_0 - n\delta_{\frac{1}{n^2}})(x) \right| \\ &= \sup_{y \in [0, 1]} \left| ny - n \left| \frac{1}{n^2} - y \right| \right| \end{aligned}$$

Let $y \in [0, 1]$. If $y \leq 1/n^2$ then

$$\left| ny - n \left| \frac{1}{n^2} - y \right| \right| = n \cdot \left| 2y - \frac{1}{n^2} \right| \leq 2ny + \frac{1}{n} \leq \frac{3}{n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Otherwise, $y > 1/n^2$ and

$$\left| ny - n \left| \frac{1}{n^2} - y \right| \right| = n \cdot \left| \frac{1}{n^2} \right| = \frac{1}{n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

It follows that $\|\phi(\mu_n) - \phi(\delta_0)\| \rightarrow 0$; that is, $\phi(\mu_n) \rightarrow \phi(\delta_0)$ in $C([0, 1])$.

Now, let $f: [0, 1] \rightarrow \mathbb{R}$ be the function defined for $x \in [0, 1]$ by $f(x) = x^{1/2}$. Then $f \in C([0, 1])$ and

$$\int f d\mu_n = (n+1)f(0) - nf\left(\frac{1}{n^2}\right) = -1,$$

but

$$\int f d\delta_0 = f(0) = 0.$$

Hence $\mu_n \not\rightarrow^* \delta_0$ as $n \rightarrow \infty$ and ϕ^{-1} is not continuous at $\phi(\delta_0)$. By linearity properties of the integral, ϕ^{-1} is discontinuous with respect to the weak-* topology on $\mathcal{M}([0, 1])$ at each point if and only if it is discontinuous at one point. The result follows. \square

Note that it is a consequence of Part 5 of the previous result that ϕ^{-1} is discontinuous at each point with respect to the norm topology on $\mathcal{M}([0, 1])$.

Corollary 3.2.18. *$I: \mathcal{M}([0, 1])^2 \rightarrow \mathbb{R}$ and $I: \mathcal{M}([0, 1]) \rightarrow \mathbb{R}$ are discontinuous at each point.*

Proof. The result follows from Corollary 3.2.13 and Corollary 3.2.14 and Theorem 3.2.17. \square

3.3 Summary

This chapter introduced the functions $d_\mu: X \rightarrow \mathbb{R}$ and $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$, where X is a compact metric space. An analysis of the functions d_μ was simple; it was easily shown that $d_\mu \in C(X)$ for each $\mu \in \mathcal{M}(X)$, and that if $(\mu_\alpha) \in \mathcal{M}(X)$ is a convergent net with limit μ such that $(\|\mu_\alpha\|)$ is bounded, then (d_{μ_α}) converges uniformly to d_μ .

An analysis of the functional $I: \mathcal{M}(X)^2 \rightarrow \mathbb{R}$ was more difficult, however, it was possible to demonstrate in general that this function is bounded and continuous on certain subsets of $(\mathcal{M}(X))^2$, is separately continuous, is continuous at each point if and only if it is continuous at one point, and is discontinuous at each

point when $\{d_\mu: \mu \in \mathcal{M}(X)\}$ is dense in $C(X)$. The chapter concluded by showing this density property in the case that $X = [0, 1]$, from which it follows that both $I: \mathcal{M}([0, 1])^2 \rightarrow \mathbb{R}$ and $I: \mathcal{M}([0, 1]) \rightarrow \mathbb{R}$ are discontinuous at each point. Further properties of the functional I may be found in a paper of Nickolas and Wolf [40].

Chapter 4

Results in compact spaces

Chapter 1 introduced the average distances $m(X)$, $M(X)$ and $\overline{M}(X)$ defined for a compact (connected) metric space X . Using the measure theory developed in Chapters 2 and 3, we will now prove the Gross-Stadje Theorem, and develop an additional characterisation of each average distance. We will also discuss properties of $m(X)$, $M(X)$ and $\overline{M}(X)$, and demonstrate various techniques to calculate their values for certain concrete spaces X .

4.1 The Gross-Stadje Theorem

Chapter 1 introduced the Gross-Stadje Theorem and showed by a direct argument that $m([0, 1])$ is defined and equal to $1/2$. We will now prove the theorem for a general compact connected metric space, using a proof which depends on Bauer's Maximum Principle and Ville's Minimax Theorem. We shall state these well-known results without proof, as to do so would require use of techniques from convexity theory and game theory which are not considered in this thesis.

Choquet discusses Bauer's Maximum Principle in [14, pg. 102].

Theorem 4.1.1 (Bauer's Maximum Principle). *Let X be a convex compact subset of a topological vector space and let $f: X \rightarrow \mathbb{R}$ be a convex continuous function. Then there exists an extreme point of X at which f attains its maximum value.*

Nikaidô discusses Ville's Minimax Theorem in [42, pg. 69].

Theorem 4.1.2 (Ville's Minimax Theorem). *Let X be a compact subset of a topological vector space and let $f: X \times X \rightarrow \mathbb{R}$ be a continuous function with respect to the product topology on $X \times X$. Then*

$$\max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} \iint f(x, y) d\mu(x) d\nu(y) = \min_{\nu \in \mathcal{M}^1(X)} \max_{\mu \in \mathcal{M}^1(X)} \iint f(x, y) d\mu(x) d\nu(y).$$

The following proof of the Gross-Stadje Theorem has been constructed from several sources. The existence proof is due to Yost [57, pg. 332-333], and is of interest as it is an elementary proof which does not require game theory, as in the proof due to Gross [24, pg. 50-52], nor measure theory and integration, as in the proof due to Stadje [48, pg. 276-277]. The uniqueness proof, which uses the above two theorems, is due to Cleary et al. [16, pg. 268]. The formula for the Gross-Stadje number $m(X)$ which uses the functional $I(\cdot, \cdot)$ is due to Stadje.

Lemma 4.1.3. *Let X and Y be Hausdorff spaces such that X is dense in Y and Y is complete, let $f: X \rightarrow \mathbb{R}$ be a continuous function and let $g: Y \rightarrow \mathbb{R}$ be the function defined for $x \in Y$ by*

$$g(x) = \begin{cases} f(x) & \text{if } x \in X, \\ \lim_{n \rightarrow \infty} f(x_n) & \text{if } x \notin X, \end{cases}$$

where for each $x \in Y \setminus X$, $(x_n) \in X$ is a convergent sequence such that $x_n \rightarrow x$. Then g is a well-defined continuous extension of f .

Proof. Let $x \in X$, let $(x_n), (y_n) \in X$ be convergent sequences with limit x and let $(z_n) \in X$ be the sequence $x_1, y_1, x_2, y_2, \dots$. Then $z_n \rightarrow x$, and each of $(x_n), (y_n)$ and (z_n) are Cauchy. Since f is continuous, it follows that each of $(f(x_n)), (f(y_n))$ and $(f(z_n))$ are also Cauchy. Hence, given that Y is complete, there exist $u, v, w \in Y$ such that $f(x_n) \rightarrow u$ and $f(y_n) \rightarrow v$ and $f(z_n) \rightarrow w$. As $(f(x_n))$ and $(f(y_n))$ are subsequences of $(f(z_n))$, it follows that $w = u = v$. Therefore

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} f(y_n),$$

giving that g is well-defined.

It is clear that g is an extension of f . It remains to be shown that g is continuous. Let d be the metric of Y , let $y_0 \in Y$ and let $(y_n) \in Y$ be a sequence such that $y_n \rightarrow y_0$. For each $i = 0, 1, \dots$, let $(x_{i,n}) \in X$ be a sequence such that $x_{i,n} \rightarrow y_i$. Then by the definition of g , $g(x_{i,n}) \rightarrow g(y_i)$ for each i . Now, for all $\varepsilon_1 > 0$ there exists n_1 such that $n > n_1 \Rightarrow d(y_n, y_0) < \varepsilon/2$, and for each $n > n_1$ there exists $m_1(n)$ such that $m > m_1(n) \Rightarrow d(x_{n,m}, y_n) < \varepsilon/2$. Hence, if $n > n_1$ and $m > m_1(n)$ then by the triangle inequality

$$d(x_{n,m}, y_0) \leq d(x_{n,m}, y_n) + d(y_n, y_0) < \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{2} = \varepsilon_1.$$

That is, $x_{n,m} \rightarrow y_0$ as $n, m \rightarrow \infty$, giving that $g(x_{n,m}) \rightarrow g(y_0)$ as $n, m \rightarrow \infty$.

Let $\varepsilon_2 > 0$. Then there exists n_2 such that for each $n > n_2$, there exists $m_2(n)$ such that for each $n > n_2$ and each $m > m_2(n)$, $|g(x_{n,m}) - g(y_0)| < \varepsilon_2/2$. Also, for each $n > n_2$, there exists $m_3(n)$ such that for each $n > n_2$ and each $m > m_3(n)$, $|g(y_n) - g(x_{n,m})| < \varepsilon_2/2$. Now, let $m_4 = \max\{m_2, m_3\}$. Then for each $n > n_2$ and each $m > m_4(n)$

$$|g(y_n) - g(y_0)| \leq |g(y_n) - g(x_{n,m})| + |g(x_{n,m}) - g(y_0)| \leq \frac{\varepsilon_2}{2} + \frac{\varepsilon_2}{2} = \varepsilon.$$

Hence, g is continuous. □

Theorem 4.1.4 (Gross-Stadje Theorem). *Let (X, d) be a compact connected metric space. Then there exists a unique $m(X, d) \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ and for all $x_1, \dots, x_n \in X$ there exists $y \in X$ such that*

$$\frac{1}{n} \sum_{i=1}^n d(x_i, y) = m(X, d).$$

Further,

$$m(X, d) = \max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} I(\mu, \nu) = \max_{\mu \in \mathcal{M}^1(X)} \min_{x \in X} d_\mu(x) = \min_{\mu \in \mathcal{M}^1(X)} \max_{x \in X} d_\mu(x).$$

Proof. Let $\mathcal{F} = \bigcup_{n=1}^{\infty} X^n$. For each $F = (x_1, \dots, x_n) \in \mathcal{F}$, let $d_F: X \rightarrow \mathbb{R}$ be the function defined for $x \in X$ by

$$d_F(x) = \frac{1}{n} \sum_{i=1}^n d(x, x_i).$$

Each d_F , being the sum of continuous functions, is continuous. Let $a: \mathcal{F} \rightarrow \mathbb{R}$ and $b: \mathcal{F} \rightarrow \mathbb{R}$ be the functions defined for $F \in \mathcal{F}$ by

$$a(F) = \min_{x \in X} d_F(x) \quad \text{and} \quad b(F) = \max_{x \in X} d_F(x).$$

As X is compact, the above values are well-defined. Recall that the continuous image of a compact connected set is compact connected, and that the compact connected subsets of \mathbb{R} are the closed bounded intervals. It follows that for each $F \in \mathcal{F}$, $d_F(X) = [a(F), b(F)]$. To prove the existence of such an $m(X, d) \in \mathbb{R}$, it will be sufficient to show that

$$\bigcap_{F \in \mathcal{F}} d_F(X) \neq \emptyset.$$

Let $F = (x_1, \dots, x_m) \in \mathcal{F}$ and let $G = (y_1, \dots, y_n) \in \mathcal{F}$. As a metric is symmetric, we have that

$$\begin{aligned} m \sum_{i=1}^n d_F(y_i) &= m \sum_{i=1}^n \frac{1}{m} \sum_{j=1}^m d(y_i, x_j) \\ &= \sum_{j=1}^m \sum_{i=1}^n d(x_j, y_i) \\ &= n \sum_{j=1}^m \frac{1}{n} \sum_{i=1}^n d(x_j, y_i) \\ &= n \sum_{j=1}^m d_G(x_j). \end{aligned}$$

Suppose that $d_F(y_i) > d_G(x_j)$ for all i and j . Then

$$\sum_{i=1}^n \sum_{j=1}^m d_F(y_i) > \sum_{i=1}^n \sum_{j=1}^m d_G(x_j),$$

giving the contradiction

$$m \sum_{i=1}^n d_F(y_i) > n \sum_{j=1}^m d_G(x_j).$$

Hence $d_F(y_i) \leq d_G(x_j)$ for some i and j , giving that $a(F) \leq b(G)$. As this holds for all $F, G \in \mathcal{F}$, we obtain

$$\sup_{F \in \mathcal{F}} a(F) \leq \inf_{F \in \mathcal{F}} b(F).$$

It follows that $\bigcap_F d_F(X) \neq \emptyset$, establishing the existence of such an $m(X, d) \in \mathbb{R}$.

As each

$$m \in \left[\sup_{F \in \mathcal{F}} a(F), \inf_{F \in \mathcal{F}} b(F) \right]$$

has the property required of $m(X, d)$, to demonstrate that such a number is unique, it remains to be shown that

$$\sup_{F \in \mathcal{F}} a(F) = \inf_{F \in \mathcal{F}} b(F) = \max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} I(\mu, \nu).$$

Note that by Corollary 3.2.2 and the compactness of $\mathcal{M}^1(X)$, the above values are well-defined.

Now, each member of \mathcal{F} can be naturally associated with a member of $\mathcal{F}^1(X)$, the set of atomic probability measures, by way of the mapping $(x_1, \dots, x_n) \mapsto \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$. We now regard a and b as the functions $a: \mathcal{F}^1(X) \rightarrow \mathbb{R}$ and $b: \mathcal{F}^1(X) \rightarrow \mathbb{R}$ defined for $\mu \in \mathcal{F}^1(X)$ by

$$a(\mu) = \min_{x \in X} d_\mu(x) \quad \text{and} \quad b(\mu) = \max_{x \in X} d_\mu(x).$$

Let $\mu \in \mathcal{F}^1(X)$ and let $(\mu_n) \in \mathcal{F}^1(X)$ be a sequence such that $\mu_n \rightarrow^* \mu$. Then by Lemma 3.1.3, (d_{μ_n}) converges uniformly to d_μ and it is a routine exercise to show that $a(\mu_n) \rightarrow a(\mu)$ and $b(\mu_n) \rightarrow b(\mu)$. We then have that a and b are continuous. Now, let $\nu \in \mathcal{M}^1(X)$ and let $(\nu_n) \in \mathcal{F}^1(X)$ be a sequence such that $\nu_n \rightarrow^* \nu$. As $\mathcal{F}^1(X)$ is dense in the complete space $\mathcal{M}^1(X)$, Lemma 4.1.3 asserts that the functions $\alpha: \mathcal{M}^1(X) \rightarrow \mathbb{R}$ and $\beta: \mathcal{M}^1(X) \rightarrow \mathbb{R}$ defined by

$$\alpha(\nu) = \lim_{n \rightarrow \infty} a(\nu_n) \quad \text{and} \quad \beta(\nu) = \lim_{n \rightarrow \infty} b(\nu_n),$$

are continuous extensions of a and b . By the definition of $\alpha(\nu)$ and the fact that $d_{\nu_n} \rightarrow d_\nu$, we have that

$$\min_{x \in X} d_{\nu_n}(x) \rightarrow \alpha(\nu) \quad \text{and} \quad \min_{x \in X} d_{\nu_n}(x) \rightarrow \min_{x \in X} d_\nu(x),$$

That is,

$$\alpha(\nu) = \min_{x \in X} d_\nu(x),$$

and similarly,

$$\beta(\nu) = \max_{x \in X} d_\nu(x).$$

Also, it is easily shown that

$$\begin{aligned} \sup_{F \in \mathcal{F}} a(F) &= \max_{\mu \in \mathcal{M}^1(X)} \alpha(\mu) = \max_{\mu \in \mathcal{M}^1(X)} \min_{x \in X} d_\mu(x) \\ \inf_{F \in \mathcal{F}} b(F) &= \min_{\mu \in \mathcal{M}^1(X)} \beta(\mu) = \min_{\mu \in \mathcal{M}^1(X)} \max_{x \in X} d_\mu(x). \end{aligned}$$

Let $\nu \in \mathcal{M}^1(X)$. Using Proposition 3.2.4, the mapping $\mathcal{M}^1(X) \rightarrow \mathbb{R}$ such that $\mu \mapsto I(\mu, \nu)$ is continuous and linear, and hence convex. Bauer's maximum principle then gives that this mapping will attain its maximum value at some extreme point of $\mathcal{M}^1(X)$, which by Proposition 2.1.2 is some measure δ_{x_0} , where $x_0 \in X$. Hence

$$\begin{aligned} \max_{\mu \in \mathcal{M}^1(X)} I(\mu, \nu) &= \max_{x \in X} I(\delta_x, \nu) \\ &= \max_{x \in X} \iint f(z, y) d\delta_x(z) d\nu(y) \\ &= \max_{x \in X} \int f(x, y) d\nu(y) \\ &= \max_{x \in X} d_\nu(x) \\ &= b(\nu), \end{aligned}$$

and by taking the minimum over $\nu \in \mathcal{M}^1(X)$, we obtain

$$\min_{\nu \in \mathcal{M}^1(X)} \max_{\mu \in \mathcal{M}^1(X)} I(\mu, \nu) = \min_{\nu \in \mathcal{M}^1(X)} b(\nu) = \inf_{F \in \mathcal{F}} b(F).$$

Similarly,

$$\max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} I(\mu, \nu) = \max_{\mu \in \mathcal{M}^1(X)} a(\mu) = \sup_{F \in \mathcal{F}} a(F).$$

Using Ville's Minimax Theorem, we have

$$\sup_{F \in \mathcal{F}} a(F) = \inf_{F \in \mathcal{F}} b(F) = \max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} I(\mu, \nu),$$

as required. □

We note that our proof of the Gross-Stadje Theorem may also be used to demonstrate Stadje's result concerning a compact connected Hausdorff space and a continuous symmetric function. It is possible to generalise the space by investigating

weak-* convergence of probability measures on compact Hausdorff spaces (the interested reader is referred to Choquet [13, pg. 217]), where it is found that the properties we required of the space of measures still hold. It is much easier to generalise the metric to a continuous symmetric function, as each inference in the proof using the metric and its associated functionals studied in Chapter 3 relied on no properties of a metric other than continuity and symmetry. A more general approach to the functionals studied in Chapter 3 would give the generalisation immediately.

We note that Cleary et al. [16, pg. 268] make the comment that “the uniqueness of $m(X, d)$ is essentially equivalent to the mini-max theorem, whereas the existence of $m(X, d)$ corresponds to the trivial mini-max inequality. This is why the existence of $m(X, d)$ is easier to prove than [its] uniqueness”.

The following generalisation of the Gross-Stadje Theorem is due to Graham Elton, and has been announced without proof in papers of Cleary et al. [16, pg. 266], Morris and Nickolas [38, pg. 463] and Yost [57, pg. 333]. The statement of the result as given here assumes that $m(X, d)$ is defined and exists for a compact connected metric space X , as in Theorem 1.1.1.

Theorem 4.1.5 (Gross-Stadje Theorem). *Let (X, d) be a compact connected metric space. Then for all $\mu \in \mathcal{M}^1(X)$ there exists $y \in X$ such that $d_\mu(y) = m(X, d)$.*

Proof. Let $\mu \in \mathcal{M}^1(X)$. If μ is an atomic measure then there exists $x_1, \dots, x_n \in X$ such that $\mu = \sum_{i=1}^n \delta_{x_i}$. Hence using Theorem 4.1.4, there exists $y \in X$ such that

$$d_\mu(y) = \int d(x, y) d\mu(x) = \frac{1}{n} \sum_{i=1}^n d(x_i, y) = m(X, d) .$$

Otherwise, let $(\mu_n) \in \mathcal{M}^1(X)$ be a sequence of atomic measures such that $\mu_n \rightarrow^* \mu$ and let $(y_n) \in X$ be a sequence such that for each n , $d_{\mu_n}(y_n) = m(X, d)$. As X is compact, there exists a subsequence (y_{k_n}) of (y_n) which converges to some $y \in X$.

Then $\mu_{k_n} \rightarrow^* \mu$ and for all $n \in \mathbb{N}$,

$$\begin{aligned}
 |m(X, d) - d_\mu(y)| &\leq |d_{\mu_{k_n}}(y_{k_n}) - d_{\mu_{k_n}}(y)| + |d_{\mu_{k_n}}(y) - d_\mu(y)| \\
 &\leq \int |d(x, y_{k_n}) - d(x, y)| d\mu_{k_n}(x) + |d_{\mu_{k_n}}(y) - d_\mu(y)| \\
 &\leq \int d(y_{k_n}, y) d\mu_{k_n}(x) + |d_{\mu_{k_n}}(y) - d_\mu(y)| \\
 &= d(y_{k_n}, y) + |d_{\mu_{k_n}}(y) - d_\mu(y)|.
 \end{aligned}$$

Now, $d(y_{k_n}, y) \rightarrow 0$ as $n \rightarrow \infty$. As the mapping $X \rightarrow \mathbb{R}$ such that $x \mapsto d(x, y)$ is continuous and bounded,

$$|d_{\mu_{k_n}}(y) - d_\mu(y)| = \left| \int d(x, y) d\mu_{k_n}(x) - \int d(x, y) d\mu(x) \right| \rightarrow 0$$

as $n \rightarrow \infty$. It follows that $d_\mu(y) = m(X, d)$. □

4.2 Properties of $m(X, d)$

We now present some elementary properties of the Gross-Stadje number. These properties concern the permissible values for $m(X, d)$, and certain attributes of m when considered as a real-valued mapping.

The following property of $m(X, d)$ is given by both Gross [24, pg. 52] and Stadje [48, pg. 277-278]. Our proof is due to Stadje.

Proposition 4.2.1. *Let (X, d) be a compact connected metric space with at least two points. Then*

$$\frac{D(X, d)}{2} \leq m(X, d) < D(X, d).$$

Proof. As X is compact, there exist $x_1, x_2 \in X$ such that $d(x_1, x_2) = D(X, d)$. Then for all $y \in X$, by the triangle inequality we have that

$$\frac{1}{2} \sum_{i=1}^2 d(x_i, y) = \frac{1}{2} (d(x_1, y) + d(x_2, y)) \geq \frac{d(x_1, x_2)}{2} = \frac{D(X)}{2}.$$

Hence $D(X)/2 \leq m(X, d)$.

Let $r = D(X, d)/2 > 0$. As X is compact, there exist $x_1, \dots, x_n \in X$ such that the collection of open spheres centred at the x_i of radius r covers X . Then for all $y \in X$, there exists $j \in \{1, \dots, n\}$ such that $d(x_j, y) < D(X, d)/2$ and

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n d(x_i, y) &= \frac{1}{n} \sum_{i=1, i \neq j}^n d(x_i, y) + \frac{1}{n} d(x_j, y) \\ &\leq \frac{n-1}{n} D(X, d) + \frac{D(X, d)}{2n} \\ &= \frac{2n-1}{2n} D(X, d) \\ &< D(X, d). \end{aligned}$$

It follows that $m(X, d) < D(X, d)$. □

We note the following property of $m(X)$ given by both Gross [24, pg. 53] and Stadje [48, pg. 278], which is interesting for two reasons. Firstly, it shows that our bounds for $m(X)$ cannot be improved. Secondly it shows that $m(X, d)$ is not a topological property of (X, d) .

Proposition 4.2.2. *Let $D, m > 0$ such that $D/2 \leq m < D$. Then there exists a compact connected metric space (X, d) homeomorphic to the Euclidean space $[0, 1]$ (equipped with its usual metric) such that $m(X, d) = m$.*

The proof of the following property is due to Yost [57, pg. 334].

Proposition 4.2.3. *Let (X, d) be a compact connected metric space and let $\alpha \leq \beta$. If for all $\mu \in \mathcal{M}^1(X)$ there exists $y \in X$ such that $\alpha \leq d_\mu(y) \leq \beta$ then*

$$\alpha \leq m(X) \leq \beta.$$

Proof. Suppose that for all $\mu \in \mathcal{M}^1(X)$ there exists $y \in X$ such that $\alpha \leq d_\mu(y) \leq \beta$, which gives that

$$\alpha \leq \max_{y \in X} d_\mu(y) \quad \text{and} \quad \min_{y \in X} d_\mu(y) \leq \beta.$$

Hence

$$\alpha \leq \min_{\mu \in \mathcal{M}^1(X)} \max_{y \in X} d_\mu(y) = m(X) = \max_{\mu \in \mathcal{M}^1(X)} \min_{y \in X} d_\mu(y) \leq \beta. \quad \square$$

Corollary 4.2.4. *Let (X, d) be a compact connected metric space and let $\alpha \leq \beta$. If for all $\mu \in \mathcal{M}^1(X)$ of the form $\mu = (1/n) \sum_{i=1}^n \delta_{x_i}$ there exists $y \in X$ such that $\alpha \leq d_\mu(y) \leq \beta$ then $\alpha \leq m(X) \leq \beta$.*

For our remaining results, we will consider $m(X, d)$ to be a real-valued function. To do so however, we must first construct a “universe” of compact connected metric spaces. Let (S, d) be a metric space. Then $H(S, d)$ will denote the set of compact non-empty subspaces of (S, d) , and $\overline{H}(S, d)$ will denote the subset of $H(S, d)$ consisting precisely of those spaces which are connected. Note that these sets are non-empty, as they necessarily contain each singleton subspace of (S, d) . We now consider m to be a function $m: \overline{H}(S, d) \rightarrow \mathbb{R}$ which assigns to each $X \in \overline{H}(S, d)$ its Gross Stadge number.

We postpone the proof of the following result until Proposition 4.5.9, where we will show the result by directly calculating the Gross-Stadge numbers of certain subspaces of \mathbb{R}^2 .

Proposition 4.2.5. *In general, m is not monotonic with respect to subset inclusion; that is, there exists a metric space (S, d) and $A, B \in \overline{H}(S, d)$ such that $A \subseteq B$ and $m(A, d) > m(B, d)$.*

Our remaining property of the Gross-Stadge number concerns the continuity of $m: \overline{H}(S, d) \rightarrow \mathbb{R}$, a discussion of which requires a suitable metric on the space $\overline{H}(S, d)$. Let $h: H(S, d) \times H(S, d) \rightarrow \mathbb{R}$ be the function defined for $X, Y \in H(S, d)$ by

$$h(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y) \right\}.$$

Then h is the well-known *Hausdorff metric* on $H(S, d)$, a full discussion of which is found in [25, Section 28]. We will assume from now on that $H(S, d)$ and $\overline{H}(S, d)$ are equipped with this metric.

Theorem 4.2.6. *Let (S, d) be a metric space and let h denote the Hausdorff metric on $H(S, d)$. Then for all $X, Y \in H(S, d)$, $|m(Y, d) - m(X, d)| < 2h(X, Y)$. Further, $m: \overline{H}(S, d) \rightarrow \mathbb{R}$ is uniformly continuous.*

Proof. Let $X, Y \in \overline{H}(S, d)$ and let $x_1, \dots, x_n \in X$. Then there exist $y_1, \dots, y_n \in Y$ such that $d(x_i, y_i) \leq h(X, Y)$ for each i . Let $\mu = (1/n) \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}^1(X)$ and let $\nu = (1/n) \sum_{i=1}^n \delta_{y_i} \in \mathcal{M}^1(Y)$, let $y \in Y$ such that $d_\nu(y) = m(Y, d)$ and let $x \in X$ such that $d_\mu(x) \leq h(X, Y)$. Then

$$\begin{aligned}
|m(Y) - d_\mu(x)| &= \left| \frac{1}{n} \sum_{i=1}^n d(y, y_i) - \frac{1}{n} \sum_{i=1}^n d(x, x_i) \right| \\
&\leq \frac{1}{n} \sum_{i=1}^n |d(y, y_i) - d(x, x_i)| \\
&\leq \frac{1}{n} \sum_{i=1}^n |d(y, y_i) - d(x_i, y)| + \frac{1}{n} \sum_{i=1}^n |d(x_i, y) - d(x, x_i)| \\
&\leq \frac{1}{n} \sum_{i=1}^n d(x_i, y_i) + \frac{1}{n} \sum_{i=1}^n d(x, y) \\
&\leq \frac{1}{n} \sum_{i=1}^n h(X, Y) + \frac{1}{n} \sum_{i=1}^n h(X, Y) \\
&= 2h(X, Y).
\end{aligned}$$

That is, for all $\mu \in \mathcal{M}^1(X)$ of the form $\mu = (1/n) \sum_{i=1}^n \delta_{x_i}$, there exists $x \in X$ such that $m(Y, d) - 2h(X, Y) \leq d_\mu(x) \leq m(Y, d) + 2h(X, Y)$. It follows from Corollary 4.2.4 that

$$m(Y, d) - 2h(X, Y) \leq m(X, d) \leq m(Y, d) + 2h(X, Y),$$

and similarly it can be shown that

$$m(X, d) - 2h(X, Y) \leq m(Y, d) \leq m(X, d) + 2h(X, Y).$$

Hence $|m(Y, d) - m(X, d)| < 2h(X, Y)$, and m is uniformly continuous. \square

4.3 Properties of $M(X, d)$

We now present some elementary properties of $M(X, d)$. We firstly characterise $M(X, d)$ using measures, and present some results arising from this characterisation. We then consider properties of $M(X, d)$ which are similar in character to those which we discussed for $m(X, d)$.

Theorem 4.3.1. *Let (X, d) be a compact metric space. Then*

$$M(X, d) = \sup_{\mu \in \mathcal{M}^1(X)} I(\mu).$$

Proof. Let

$$A = \sup \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the supremum is taken over $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$ and $w_1, \dots, w_n > 0$ such that $\sum_{i=1}^n w_i = 1$, and let

$$B = \sup_{\mu \in \mathcal{M}^1(X)} I(\mu).$$

We want to show that $A = B$.

Let $n \in \mathbb{N}$, let $x_1, \dots, x_n \in X$ and let $w_1, \dots, w_n > 0$ such that $\sum_{i=1}^n w_i = 1$. Let $\mu = \sum_{i=1}^n w_i \delta_{x_i} \in \mathcal{M}^1(X)$. Then

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) = \iint d(x, y) d\mu(x) d\mu(y) = I(\mu) \leq B,$$

giving that $A \leq B$. It remains to be shown that $B \leq A$.

Let $\mu \in \mathcal{M}^1(X)$ and let $(\mu_n) \in \mathcal{M}^1(X)$ be a sequence of atomic measures such that $\mu_n \rightarrow^* \mu$. Then each μ_n is of the form $\sum_{i=1}^m w_i \delta_{x_i}$, where $m \in \mathbb{N}$ and $x_1, \dots, x_m \in X$ and $w_1, \dots, w_m > 0$ such that $\sum_{i=1}^m w_i = 1$. Hence $I(\mu_n)$ is of the form

$$\sum_{i=1}^m \sum_{j=1}^m w_i w_j d(x_i, x_j),$$

and $I(\mu_n) \leq A$ for each n . Since by Corollary 3.2.3 we know that I is continuous on $\mathcal{M}^1(X)$, we have that $I(\mu) \leq A$. Hence $B \leq A$, as required. \square

Given that for a compact metric space X , $\mathcal{M}^1(X)$ is compact and metrizable and $I: \mathcal{M}^1(X) \rightarrow \mathbb{R}$ is continuous, it is a standard compactness argument to show that $M(X) = I(\mu)$ for some $\mu \in \mathcal{M}^1(X)$. We say that μ is an *M-maximal measure*.

Definition 4.3.2. Let X be a compact metric space and let $\mu \in \mathcal{M}^1(X)$. If $M(X) = I(\mu)$ then μ is an *M-maximal measure*.

The following is found in Björck [8, pg. 256].

Theorem 4.3.3. *There exists an M -maximal measure on every compact metric space.*

Proof. Let X be a compact metric space and let $(\mu_n) \in \mathcal{M}^1(X)$ be a sequence such that $I(\mu_n) \rightarrow M(X)$. As $\mathcal{M}^1(X)$ is compact and metrizable there exists a subsequence (μ_{k_n}) and $\mu \in \mathcal{M}^1(X)$ such that $\mu_{k_n} \rightarrow^* \mu$. Then $I(\mu_{k_n}) \rightarrow M(X)$, and as by Corollary 3.2.3 we know that $I: \mathcal{M}^1(X) \rightarrow \mathbb{R}$ is continuous, it follows that $I(\mu_{k_n}) \rightarrow I(\mu)$. Hence $M(X) = I(\mu)$, giving then that μ is M -maximal. \square

We will discuss the uniqueness of M -maximal measures on compact metric spaces in Chapter 6.

The following results are due to Björck [8, pg. 256-257].

Lemma 4.3.4. *Let X be a compact metric space, let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure and let $\nu \in \mathcal{M}(X)$ such that $\nu(X) = 0$ and $\mu + \nu \geq 0$. Then $I(\mu, \nu) \leq 0$.*

Proof. Let $\varepsilon \in (0, 1]$ and let $\lambda = \mu + \varepsilon \cdot \nu$. We want to show that $\lambda \in \mathcal{M}^1(X)$. Let $A \in \mathcal{B}(X)$. If $\nu(A) \geq 0$, then it is clear that $\lambda(A) \geq 0$. Otherwise, if $\nu(A) < 0$ then $\nu(A) < \varepsilon \cdot \nu(A)$, giving that $0 \leq \mu(A) + \nu(A) < \mu(A) + \varepsilon \cdot \nu(A) = \lambda(A)$. Hence λ is a positive measure, and as $\lambda(X) = 1$ it follows that $\lambda \in \mathcal{M}^1(X)$.

Now, $M(X) \geq I(\lambda) = I(\mu) + \varepsilon^2 I(\nu) + 2\varepsilon I(\mu, \nu) = M(X) + \varepsilon^2 I(\nu) + 2\varepsilon I(\mu, \nu)$, giving that

$$I(\mu, \nu) \leq -\varepsilon \frac{I(\nu)}{2} \leq \varepsilon \left| \frac{I(\nu)}{2} \right|.$$

Letting $\varepsilon \rightarrow 0$, we obtain $I(\mu, \nu) \leq 0$. \square

Theorem 4.3.5. *Let (X, d) be a compact metric space and let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure. Then for all $x \in X$, $d_\mu(x) \leq M(X, d)$, with equality holding when x belongs to the support of μ .*

Proof. Recall that C_μ denotes the support of μ . Suppose that there exists $x_1 \in C_\mu$ and $x_2 \in X$ such that $d_\mu(x_1) < d_\mu(x_2)$. Let $a = d_\mu(x_2)$ and let $b = d_\mu(x_1)$. Then as

d_μ is continuous, there exists $\delta > 0$ such that for all $y_1, y_2 \in X$,

$$d(y_1, y_2) < \delta \Rightarrow |d_\mu(y_1) - d_\mu(y_2)| < \frac{a-b}{2}.$$

Note that $x_1 \neq x_2$. Let $r = \min\{|x_1 - x_2|, \delta\} > 0$, and let U be the open ball centred at x_1 of radius r in X . Then $x_2 \notin U$ and for all $x \in U$,

$$0 \leq d_\mu(x) \leq d_\mu(x_1) + \frac{a-b}{2} = \frac{a+b}{2}.$$

Also, $\mu(U) > 0$ as $x_1 \in C_\mu$.

Let $\nu \in \mathcal{M}(X)$ be the measure defined for $A \in \mathcal{B}(X)$ by

$$\nu(A) = \mu(U) \cdot \delta_{x_2}(A) - \mu(A \cap U).$$

Then $\nu(X) = 0$, and we have that

$$\begin{aligned} I(\mu, \nu) &= \int d_\mu d\nu \\ &= d_\mu(x_2) \cdot \mu(U) - \int_U d_\mu d\mu \\ &\geq d_\mu(x_2) \cdot \mu(U) - \int_U \frac{a+b}{2} d\mu \\ &= a \cdot \mu(U) - \frac{a+b}{2} \cdot \mu(U) \\ &= \frac{a-b}{2} \cdot \mu(U) \\ &> 0. \end{aligned}$$

Now, for all $A \in \mathcal{B}(X)$

$$\begin{aligned} &(\mu + \nu)(A) \\ &= \mu(A) + \nu(A \cap U) + \nu(A \cap (X \setminus U)) \\ &= \mu(A) + \mu(U) \cdot \delta_{x_2}(A \cap U) - \mu(A \cap U) + \mu(U) \cdot \delta_{x_2}(A \cap (X \setminus U)) - \mu(\emptyset) \\ &= \mu(A \cap (X \setminus U)) + \mu(U) \cdot (\delta_{x_2}(A \cap U) + \delta_{x_2}(A \cap (X \setminus U))) \\ &\geq 0. \end{aligned}$$

Hence by Lemma 4.3.4, $I(\mu, \nu) \leq 0$, which is a contradiction. It follows that for all $x_1 \in C_\mu$ and for all $x_2 \in X$, $d_\mu(x_1) \geq d_\mu(x_2)$. It is easily seen that d_μ is constant

on C_μ . Let $M \in \mathbb{R}$ such that for all $x \in X$, $d_\mu(x) \leq M$ with equality holding when $x \in C_\mu$. Then

$$M(X) = I(\mu) = \int_{C_\mu} d_\mu d\mu = \int_{C_\mu} M d\mu = M. \quad \square$$

We note that the previous result gives that when $\mu \in \mathcal{M}^1(X)$ is an M -maximal measure on a compact metric space X , d_μ is constant on the support of μ . Also, in general it is not required that $d_\mu(x) < M(X)$ for all $x \in X \setminus C_\mu$. For example, let $X = [0, 1]$ and let $\mu = (1/2)\delta_0 + (1/2)\delta_1 \in \mathcal{M}^1([0, 1])$. Then $I(\mu) = 1/2$ and by Proposition 1.2.3, μ is an M -maximal measure on $[0, 1]$. Note that for all $x \in C_\mu$,

$$d_\mu(x) = \frac{1}{2} \cdot |0 - x| + \frac{1}{2} \cdot |1 - x| = \frac{1}{2} = M([0, 1]).$$

We now discuss the range of permissible values for $M(X, d)$. The fact that $M(X, d) < D(X)$ was first established by Peter Nickolas, using a graph-theoretic proof, but we will prove the result using measures.

Proposition 4.3.6. *Let (X, d) be a compact metric space with at least two points. Then*

$$\frac{D(X, d)}{2} \leq M(X, d) < D(X, d).$$

Further, if (X, d) is connected then $m(X, d) \leq M(X, d)$.

Proof. As (X, d) is compact, there exist $x_1, x_2 \in X$ such that $d(x_1, x_2) = D(X, d)$.

By letting

$$\mu = \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2} \in \mathcal{M}^1(X),$$

we then obtain

$$M(X, d) \geq I(\mu) = \sum_{i=1}^2 \sum_{j=1}^2 d(x_i, x_j) = \frac{D(X, d)}{2}.$$

Now, let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure and recall that C_μ denotes the support of μ . Let $x_0 \in C_\mu$ and let $U = \{x \in X : d(x, x_0) < D(X, d)/4\}$. As d is

continuous, U is an open subset of X and since $x_0 \in C_\mu$, $\mu(U) > 0$. It follows that

$$\begin{aligned}
 d_\mu(x_0) &= \int_U d(x, x_0) d\mu(x) + \int_{X \setminus U} d(x, x_0) d\mu(x) \\
 &\leq \int_U \frac{D(X, d)}{4} d\mu + \int_{X \setminus U} D(X, d) d\mu \\
 &= D(X, d) \left(\frac{1}{4} \cdot \mu(U) + \mu(X \setminus U) \right) \\
 &< D(X, d) \cdot \mu(X) \\
 &= D(X, d).
 \end{aligned}$$

Using Theorem 4.3.5, we then have that $d_\mu(x) < D(X)$ for all $x \in X$. As the continuous image of a compact set is compact, $d_\mu(X)$ must be compact, and there exists $c \in \mathbb{R}$ such that $d_\mu(x) \leq c < D(X)$ for all $x \in X$. Hence

$$M(X, d) = I(\mu) = \int d_\mu d\mu \leq \int c d\mu = c < D(X, d).$$

Now, suppose that (X, d) is connected. Then by Theorem 4.1.4, there exists $\mu \in \mathcal{M}^1(X)$ such that

$$m(X, d) = \min_{y \in X} d_\mu(y).$$

Hence

$$m(X, d) = \int m(X, d) d\mu \leq \int d_\mu d\mu = I(\mu) \leq M(X, d). \quad \square$$

Finally, we now consider M to be a mapping $M: H(S, d) \rightarrow \mathbb{R}$ for some metric space (S, d) . Recall from Section 4.2 that $H(S, d)$ is the set of compact subspaces of (S, d) , and is itself a metric space, equipped with the Hausdorff metric.

Theorem 4.3.7. *Let (S, d) be a metric space. Then*

1. *The mapping $M: H(S, d) \rightarrow \mathbb{R}$ is monotonic with respect to subset inclusion.*
2. *For all $X, Y \in H(S, d)$, $|M(X) - M(Y)| \leq 2h(X, Y)$.*
3. *The mapping $M: H(S, d) \rightarrow \mathbb{R}$ is continuous.*

Proof. To show that M is monotonic, let $X, Y \in H(S, d)$ such that $X \subseteq Y$ and let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure on X . Noting that $X \in \mathcal{B}(Y)$, let $\nu \in \mathcal{M}^1(Y)$ be defined for $A \in \mathcal{B}(Y)$ by $\nu(A) = \mu(A \cap X)$. Then

$$M(X) = I(\mu) = I(\nu) \leq M(Y).$$

Let $X, Y \in H(S, d)$, let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure on X and let $(\mu_n) \in \mathcal{M}^1(X)$ be a sequence of atomic measures such that $\mu_n \rightarrow^* \mu$. For each n , write $\mu_n = \sum_{i=1}^{k_n} w_{ni} \delta_{x_{ni}}$ where for each $n \in \mathbb{N}$, $k_n \in \mathbb{N}$ and for each $i = 1, \dots, k_n$, $x_{ni} \in X$. Given that for each $x_{ni} \in X$, there exists $y_{ni} \in Y$ such that $d(x_{ni}, y_{ni}) \leq h(X, Y)$, let $(\nu_n) \in \mathcal{M}^1(Y)$ be a sequence defined for $n \in \mathbb{N}$ by $\nu_n = \sum_{i=1}^{k_n} w_{ni} \delta_{y_{ni}}$. We then have that

$$\begin{aligned} & |I(\mu_n) - I(\nu_n)| \\ &= \left| \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} w_{ni} w_{nj} d(x_{ni}, x_{nj}) - \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} w_{ni} w_{nj} d(y_{ni}, y_{nj}) \right| \\ &\leq \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} w_{ni} w_{nj} (|d(x_{ni}, x_{nj}) - d(x_{ni}, y_{nj})| + |d(x_{ni}, y_{nj}) - d(y_{ni}, y_{nj})|) \\ &\leq \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} w_{ni} w_{nj} (d(x_{nj}, y_{nj}) + d(x_{ni}, y_{ni})) \\ &\leq \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} w_{ni} w_{nj} \cdot 2h(X, Y) \\ &\leq 2h(X, Y) \sum_{i=1}^{k_n} w_{ni} \sum_{j=1}^{k_n} w_{nj} \\ &= 2h(X, Y). \end{aligned}$$

Hence for each n , $I(\mu_n) \leq 2h(X, Y) + I(\nu_n) \leq 2h(X, Y) + M(Y)$, and as $I(\cdot)$ is continuous on $\mathcal{M}^1(X)$, we then have that $M(X) \leq 2h(X, Y) + M(Y)$. Similarly, $M(Y) \leq 2h(X, Y) + M(X)$, giving that

$$|M(X) - M(Y)| \leq 2h(X, Y).$$

That M is continuous is immediate. □

4.4 Properties of $\overline{M}(X, d)$

We now characterise $\overline{M}(X, d)$ using measures. Aside from this characterisation, for now we will merely comment on properties of $\overline{M}(X, d)$. We will find it necessary to study *quasihypermetric spaces* to provide a meaningful discussion on $\overline{M}(X, d)$. This class of metric spaces is investigated in Chapter 5, and we will return to studying properties of $\overline{M}(X, d)$ in Chapter 6.

Proposition 4.4.1. *Let (X, d) be a compact metric space. Then*

$$\overline{M}(X, d) = \sup I(\mu),$$

where the supremum is taken over all $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$.

Proof. Let

$$A = \sup \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j),$$

where the supremum is taken over $n \in \mathbb{N}$ and $x_1, \dots, x_n \in X$ and $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 1$, and let

$$B = \sup I(\mu),$$

where the supremum is taken over $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$. We want to show that $A = B$. As in the proof of Proposition 4.3.1, it is easily shown that $A \leq B$, and it will be sufficient to show that $B \leq A$.

Let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$ and let $\varepsilon > 0$. Noting that $\|\mu\| \neq 0$, by Lemma 2.3.23 there exists a partition $A_1, \dots, A_n \in \mathcal{B}(X)$ of X such that for each i ,

$$D(A_i) < \frac{\varepsilon}{2\|\mu\|^2}.$$

Let $\nu = \sum_{i=1}^n \mu(A_i) \delta_{x_i} \in \mathcal{M}(X)$, where each x_i is some point belonging to A_i . Then $\nu(X) = 1$ and

$$I(\nu) = \sum_{i=1}^n \sum_{j=1}^n \mu(A_i) \mu(A_j) d(x_i, x_j) \leq A.$$

Note that

$$\|\nu\| \leq \sum_{i=1}^n |\mu(A_i)| \leq \sum_{i=1}^n |\mu|(A_i) = |\mu|(X) = \|\mu\| ,$$

and

$$\int_{A_i} d_\mu d\nu = \mu(A_i) d_\mu(x_i) = \int_{A_i} d_\mu(x_i) d\mu .$$

We then have

$$\begin{aligned} |I(\mu) - I(\mu, \nu)| &= \left| \sum_{i=1}^n \int_{A_i} d_\mu d\mu - \sum_{i=1}^n \int_{A_i} d_\mu d\nu \right| \\ &\leq \sum_{i=1}^n \left| \int_{A_i} d_\mu d\mu - \int_{A_i} d_\mu d\nu \right| \\ &\leq \sum_{i=1}^n \left| \int_{A_i} (d_\mu(x) - d_\mu(x_i)) d\mu(x) \right| \\ &\leq \sum_{i=1}^n \int_{A_i} |d_\mu(x) - d_\mu(x_i)| d|\mu|(x). \end{aligned}$$

Now, for each i and for all $x \in A_i$,

$$\begin{aligned} |d_\mu(x) - d_\mu(x_i)| &= \left| \int (d(x, y) - d(x_i, y)) d\mu(y) \right| \\ &\leq \int |d(x, y) - d(x_i, y)| d|\mu|(y) \\ &\leq \int d(x, x_i) d|\mu|(y) \\ &< \int \frac{\varepsilon}{2 \|\mu\|^2} d|\mu| \\ &= \frac{\varepsilon}{2 \|\mu\|} . \end{aligned}$$

It follows that

$$|I(\mu) - I(\mu, \nu)| < \sum_{i=1}^n \int_{A_i} \frac{\varepsilon}{2 \|\mu\|} d|\mu| = \frac{\varepsilon}{2 \|\mu\|} \sum_{i=1}^n |\mu|(A_i) = \frac{\varepsilon}{2 \|\mu\|} |\mu|(X) = \frac{\varepsilon}{2} .$$

Similarly, $|I(\nu, \mu) - I(\nu)| < \varepsilon/2$. Hence

$$|I(\mu) - I(\nu)| \leq |I(\mu) - I(\mu, \nu)| + |I(\nu, \mu) - I(\nu)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon ,$$

which gives that $I(\mu) < I(\nu) + \varepsilon \leq A + \varepsilon$. By taking the infimum over ε , we obtain

$I(\mu) \leq A$, which gives that $B \leq A$. \square

Given a compact metric space X , it is now natural to ask if $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ attains the value $\overline{M}(X)$ for some $\mu \in \mathcal{M}(X)$ of unit mass, as we did for $M(X)$.

Definition 4.4.2. Let X be a compact metric space and let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$. If $\overline{M}(X) = I(\mu)$ then μ is an \overline{M} -maximal measure.

We state without proof the following result of Alexander [1, pg. 317-318].

Proposition 4.4.3. Let X be the closed unit ball in \mathbb{R}^3 . Then $\overline{M}(X) = 2$ and there does not exist an \overline{M} -maximal measure on X .

The following result concerning the range of permissible values for $\overline{M}(X, d)$ may be proven using essentially the same argument which gives Propositions 4.3.6.

Proposition 4.4.4. Let X be a compact metric space. Then

$$\frac{D(X)}{2} \leq M(X) \leq \overline{M}(X) \leq \infty.$$

Further, if (X, d) is connected then $m(X, d) \leq \overline{M}(X, d)$.

Finally, as in Section 4.3, we now consider M to be a mapping $M: H(S, d) \rightarrow \mathbb{R}$. We will investigate if this mapping is continuous in Chapter 6, but for now we assert that this mapping is monotonic. The result may be proven by essentially the same argument which gives Proposition 4.3.7.

Proposition 4.4.5. For all metric spaces (S, d) , the mapping $\overline{M}: H(S, d) \rightarrow \mathbb{R}$ is monotonic with respect to subset inclusion.

4.5 Techniques for calculating $m(X)$ and $M(X)$

We will now calculate $m(X)$ and $M(X)$ for various compact (connected) metric spaces X . These constants are considerably easier to calculate in most spaces than $\overline{M}(X)$. We will discuss values of $\overline{M}(X)$ in Chapter 6. The examples we include here present a variety of techniques which may be used to calculate $m(X)$ and $M(X)$.

4.5.1 Max-min measures

Recall that if X is a compact connected metric space, then

$$m(X) = \max_{\mu \in \mathcal{M}^1(X)} \min_{y \in X} d_\mu(y) = \min_{\mu \in \mathcal{M}^1(X)} \max_{y \in X} d_\mu(y).$$

Definition 4.5.1. Let (X, d) be a compact connected metric space. If $\mu, \nu \in \mathcal{M}^1(X)$ such that

$$m(X, d) = \min_{y \in X} d_\mu(y) \quad \text{and} \quad m(X, d) = \max_{y \in X} d_\nu(y),$$

then μ is a *max-min* measure on X and ν is a *min-max* measure on X .

Note that the above formulae for the Gross-Stadje number establish the existence of max-min and min-max measures on compact connected metric spaces.

Cleary et al. [16, pg. 261-262] have used max-min and min-max measures to calculate $m(X)$, where X is the perimeter of an equilateral triangle. The proof requires a lemma of Cleary [15, pg. 44-45].

Lemma 4.5.2. *Let L be a line segment in \mathbb{R}^2 and let $a, b \in \mathbb{R}^2$ lie on a line segment parallel to L such that the midpoint of a and b lies on the perpendicular bisector of L . Then*

$$\min_{y \in L} (\|a - y\| + \|b - y\|)$$

is given by the midpoint of L , and

$$\max_{y \in L} (\|a - y\| + \|b - y\|)$$

is given by an endpoint of L .

Proof. Let $y \in L$. Then by the triangle inequality, $\|a - y\| + \|b - y\| \geq \|a - b\|$. As $y = (a + b)/2$ implies that $\|a - y\| + \|b - y\| = \|a - b\|$, we then have that

$$\min_{y \in L} (\|a - y\| + \|b - y\|)$$

is given by the midpoint of L .

Let c be an endpoint of L . Then $\|a - y\| + \|b - y\| \leq \|a - y\| + \|b - c\| + \|c - y\|$. Now, $\|a - y\| + \|c - y\| \rightarrow \|a - c\|$ as $y \rightarrow c$. Hence

$$\|a - y\| + \|b - y\| \leq \|a - c\| + \|b - c\| ,$$

giving that

$$\max_{y \in L} (\|a - y\| + \|b - y\|)$$

is given by an endpoint of L . □

Proposition 4.5.3. *Let X be the perimeter of an equilateral triangle with sides of unit length. Then*

$$m(X) = \frac{2 + \sqrt{3}}{6} .$$

Proof. Let x_1, x_2 and x_3 be the vertices of X and let

$$\mu = \frac{1}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2} + \frac{1}{3}\delta_{x_3} \in \mathcal{M}^1(X) .$$

We want to minimise d_μ on X . Now for all $y \in X$, $d_\mu(y) = (1/3) \sum_{i=1}^3 \|x_i - y\|$, and by symmetry d_μ will attain its minimum value for some y belonging to the line segment joining x_1 to x_2 . By Lemma 4.5.2, the value of

$$\min_{y \in \overline{x_1 x_2}} d_\mu(y) = (1/3) (1 + \|x_3 - y\|)$$

occurs when y is the midpoint of $\overline{x_1 x_2}$. Hence

$$\min_{y \in X} d_\mu(y) = \frac{1}{3} \left(1 + \frac{\sqrt{3}}{2} \right) = \frac{2 + \sqrt{3}}{6} .$$

Let x'_1, x'_2 and x'_3 be the midpoints of the edges of X and let

$$\nu = \frac{1}{3}\delta_{x'_1} + \frac{1}{3}\delta_{x'_2} + \frac{1}{3}\delta_{x'_3} \in \mathcal{M}^1(X) .$$

We want to maximise d_ν on X . Now for all $y \in X$, $d_\nu(y) = (1/3) \sum_{i=1}^3 \|x'_i - y\|$, and by symmetry d_ν will attain its maximum value for some y belonging to the line segment joining x_1 to x_2 . By Lemma 4.5.2, the value of

$$\max_{y \in \overline{x_1 x_2}} d_\nu(y)$$

occurs when y is an endpoint of $\overline{x_1 x_2}$. Hence

$$\max_{y \in X} d_\nu(y) = \frac{1}{3} \left(1 + \frac{\sqrt{3}}{2} \right) = \frac{2 + \sqrt{3}}{6}.$$

As

$$\min_{y \in X} d_\mu(y) = \max_{y \in X} d_\nu(y),$$

we have that μ is a max-min measure and ν is a min-max measure. Therefore,

$$m(X) = \frac{2 + \sqrt{3}}{6}. \quad \square$$

Cleary and Morris [15, pg. 44-48] have generalised the proof of Proposition 4.5.3 to calculate the Gross-Stadje number of the perimeter of a regular polygon.

Proposition 4.5.4. *Let X be the perimeter of a regular n -gon inscribed in a circle of radius 1. Then the probability measure uniformly distributed on the vertices of X are a max-min measure on X , the probability measure uniformly distributed on the midpoints of the edges of X is a min-max measure on X and*

$$m(X) = \frac{1}{n} \sum_{i=0}^{n-1} \left(\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2i\pi}{n} - \cos \frac{2(i-1)\pi}{n} \right)^{1/2}.$$

Proof. Note that X is a compact connected metric space. Let P_0, \dots, P_{n-1} be the vertices of X and let Q_0, \dots, Q_{n-1} be the midpoints of the edges of X , where each Q_i is the midpoint of the edge joining P_i and P_{i+1} (note the special case that Q_{n-1} is the midpoint of the edge joining P_0 and P_{n-1}). Each P_i may then be represented by

$$P_i = \left(\cos \frac{2\pi i}{n}, \sin \frac{2\pi i}{n} \right).$$

Let $\mu \in \mathcal{M}^1(X)$ be the measure uniformly distributed on the P_i ; that is, let $\mu = (1/n) \sum_{i=0}^{n-1} \delta_{P_i}$. We want to show that

$$\min_{y \in X} d_\mu(y) = \frac{1}{n} \sum_{i=0}^{n-1} \|P_i - Q_0\|.$$

We consider two cases.

Case 1: n is even. By Lemma 4.5.2, we have that for each $i = 2, \dots, (n/2)$,

$$\min_{y \in \overline{P_0 P_1}} (\|P_i - y\| + \|P_{n+1-i} - y\|) = \|P_i - Q_0\| + \|P_{n+1-i} - Q_0\| .$$

As

$$\min_{y \in \overline{P_0 P_1}} (\|P_0 - y\| + \|P_1 - y\|) = \|P_0 - P_1\| = \|P_0 - Q_0\| + \|P_1 - Q_0\| ,$$

using the symmetry of X , it follows that

$$\begin{aligned} \min_{y \in X} d_\mu(y) &= \min_{y \in \overline{P_0 P_1}} \frac{1}{n} \sum_{i=0}^{n-1} \|P_i - y\| \\ &= \min_{y \in \overline{P_0 P_1}} \frac{1}{n} \left(\|P_0 - y\| + \|P_1 - y\| + \sum_{i=2}^{n/2} (\|P_i - y\| + \|P_{n+1-i} - y\|) \right) \\ &= \frac{1}{n} \left(\|P_0 - Q_0\| + \|P_1 - Q_0\| + \sum_{i=2}^{n/2} (\|P_i - Q_0\| + \|P_{n+1-i} - Q_0\|) \right) \\ &= \frac{1}{n} \sum_{i=1}^n \|P_i - Q_n\| . \end{aligned}$$

Case 2: n is odd. As $P_{(n+1)/2}$ and Q_0 belong to the perpendicular bisector of $\overline{P_0 P_1}$, it follows that

$$\min_{y \in \overline{P_0 P_1}} \|P_{(n+1)/2} - y\| = \|P_{(n+1)/2} - Q_0\| .$$

As in Case 1, we may calculate

$$\min_{y \in X} d_\mu(y) = \frac{1}{n} \sum_{i=1}^n \|P_i - Q_n\| .$$

Now, let $\nu \in \mathcal{M}^1(X)$ be the measure uniformly distributed on the Q_i ; that is, let $\nu = (1/n) \sum_{i=0}^{n-1} \delta_{Q_i}$. We want to show that

$$\max_{y \in X} d_\nu(y) = \frac{1}{n} \sum_{i=0}^{n-1} \|Q_i - P_0\| .$$

Again, we consider two cases.

Case 1: n is even. By Lemma 4.5.2, we have that for each $i = 0, \dots, (n/2) - 1$,

$$\max_{y \in \overline{P_0 P_1}} (\|Q_i - y\| + \|Q_{n-1-i} - y\|) = \|Q_i - P_0\| + \|Q_{n-1-i} - P_0\| .$$

Using the symmetry of X , it follows that

$$\begin{aligned} \max_{y \in X} d_\nu(y) &= \max_{y \in \overline{P_0 P_1}} \frac{1}{n} \sum_{i=0}^{n-1} \|Q_i - y\| \\ &= \min_{y \in \overline{P_0 P_1}} \frac{1}{n} \sum_{i=0}^{n/2-1} (\|Q_i - y\| + \|Q_{n-1-i} - y\|) \\ &= \frac{1}{n} \sum_{i=0}^{n/2-1} (\|Q_i - P_0\| + \|Q_{n-1-i} - P_0\|) \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \|Q_i - P_0\| . \end{aligned}$$

Case 2: n is odd. It is easily seen that

$$\max_{y \in \overline{P_1 P_n}} \|Q_{(n-1)/2} - y\| = \|Q_{(n-1)/2} - P_0\| ,$$

As in Case 1, we may calculate

$$\max_{y \in X} d_\nu(y) = \frac{1}{n} \sum_{i=0}^{n-1} \|Q_i - P_0\| .$$

Now, it is easily shown that for each i , $\|Q_i - P_0\| = \|P_i - Q_0\|$. Hence

$$m(X) = \max_{y \in X} d_\nu(y) = \frac{1}{n} \sum_{i=1}^n \|P_i - Q_0\| = \min_{y \in X} d_\mu(y) .$$

For each i ,

$$\begin{aligned} &\|P_i - Q_0\| \\ &= \left\| \left(\cos \frac{2\pi i}{n}, \sin \frac{2\pi i}{n} \right) - \frac{1}{2} \left(1 + \cos \frac{2\pi}{n}, \sin \frac{2\pi}{n} \right) \right\| \\ &= \left\| \left(\cos \frac{2\pi i}{n} - \frac{1}{2} - \frac{1}{2} \cos \frac{2\pi}{n}, \sin \frac{2\pi i}{n} - \frac{1}{2} \sin \frac{2\pi}{n} \right) \right\| \\ &= \left(\left(\cos \frac{2\pi i}{n} - \frac{1}{2} - \frac{1}{2} \cos \frac{2\pi}{n} \right)^2 + \left(\sin \frac{2\pi i}{n} - \frac{1}{2} \sin \frac{2\pi}{n} \right)^2 \right)^{1/2} \end{aligned}$$

$$\begin{aligned}
&= \left(1 + \frac{1}{4} + \frac{1}{4} - \cos \frac{2\pi i}{n} - \cos \frac{2\pi i}{n} \cos \frac{2\pi}{n} + \frac{1}{2} \cos \frac{2\pi}{n} - \sin \frac{2\pi i}{n} \sin \frac{2\pi}{n} \right)^{1/2} \\
&= \left(\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2\pi i}{n} - \cos \left(\frac{2\pi i}{n} - \frac{2\pi}{n} \right) \right)^{1/2} \\
&= \left(\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2\pi i}{n} - \cos \frac{2\pi(i-1)}{n} \right)^{1/2}.
\end{aligned}$$

We then have that

$$m(X) = \frac{1}{n} \sum_{k=0}^{n-1} \left(\frac{3}{2} + \frac{1}{2} \cos \frac{2\pi}{n} - \cos \frac{2k\pi}{n} - \cos \frac{2(k-1)\pi}{n} \right)^{1/2}. \quad \square$$

4.5.2 d -invariant measures

We now consider compact connected metric spaces which give rise to a probability measure which is both a max-min and a min-max measure.

Definition 4.5.5. Let (X, d) be a metric space and let $\mu \in \mathcal{M}(X)$ such that the value of $d_\mu(y)$ is independent of $y \in X$. Then μ is a d -invariant measure.

If (X, d) is a compact connected metric space, then it is clear from the strong form of the Gross-Stadje Theorem that the existence of a d -invariant measure $\mu \in \mathcal{M}^1(X)$ greatly simplifies the calculation of $m(X, d)$. In this case, we simply calculate

$$m(X, d) = d_\mu(y) = \int d(x, y) d\mu(x)$$

for some convenient $y \in X$. Note $\mu \in \mathcal{M}^1(X)$ is d -invariant if and only if μ is both a max-min and a min-max measure.

By way of example, let d be the Euclidean metric on $[0, 1]$. Then

$$\mu = (1/2)(\delta_0 + \delta_1) \in \mathcal{M}^1([0, 1])$$

is a d -invariant measure on $[0, 1]$, and

$$m([0, 1]) = \int \|x - 0\| d\mu(x) = \frac{1}{2} \|0\| + \frac{1}{2} \|1\| = \frac{1}{2}.$$

Without using the strong form of the Gross-Stadje Theorem, nor the existence of max-min and min-max measures, Morris and Nickolas [38, pg. 459-460] proved that the existence of a d -invariant measure $\mu \in \mathcal{M}^1(X)$ still allows $m(X, d)$ to be calculated in the same manner. Their proof requires Fubini's Theorem and Stadje's formula

$$m(X, d) = \max_{\mu \in \mathcal{M}^1(X)} \min_{\nu \in \mathcal{M}^1(X)} I(\mu, \nu).$$

We now use the existence of a well-known d -invariant measure to calculate the Gross-Stadje number of each $(n-1)$ -dimensional sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ in \mathbb{R}^n . Note that each such sphere is compact and connected.

Proposition 4.5.6. *Let d be the usual metric on S^1 . Then the usual normalised measure on S^1 is d -invariant and $m(S^1) = 4/\pi$.*

Proof. For all $0 \leq \theta < 2\pi$, let $P_\theta = (\cos \theta, \sin \theta)$. Then $S^1 = \{P_\theta : 0 \leq \theta < 2\pi\}$. Observe that for all $0 \leq \theta < 2\pi$,

$$\begin{aligned} \|P_\theta - P_0\| &= \|(\cos \theta - 1, \sin \theta)\| \\ &= (\cos^2 \theta - 2 \cos \theta + 1 + \sin^2 \theta)^{1/2} \\ &= (2 - 2 \cos \theta)^{1/2} \\ &= \left(4 \sin^2 \frac{\theta}{2}\right)^{1/2} \\ &= 2 \sin \frac{\theta}{2}. \end{aligned}$$

Now, let $\mu \in \mathcal{M}^1(S^1)$ denote the usual normalised measure on S^1 . Using a suitable rotation, it is clear that μ is a d -invariant measure, where d is the metric of S^1 .

Hence

$$m(S^1) = \int \|x - P_0\| d\mu(x) = \int_0^{2\pi} \left(2 \sin \frac{\theta}{2}\right) d\left(\frac{1}{2\pi}\theta\right) = \frac{4}{\pi}. \quad \square$$

Morris and Nickolas present an elementary proof in [38, pg. 461-462] of the above result, assuming only that $m(S^1)$ exists and is unique. We now generalise the above result to apply to all finite-dimensional spheres. The formula is due to Morris and Nickolas [38].

Proposition 4.5.7. *Let $n \in \mathbb{N}$ such that $n \geq 2$ and let d be the usual metric on S^{n-1} . Then the usual normalised measure on S^{n-1} is d -invariant and*

$$m(S^{n-1}) = \frac{2^{n-1} \left(\Gamma\left(\frac{n}{2}\right) \right)^2}{\sqrt{\pi} \Gamma\left(\frac{2n-1}{2}\right)},$$

where Γ denotes the gamma function.

Proof. Let d denote the metric of S^{n-1} and let $\mu \in \mathcal{M}(S^{n-1})$ denote the usual normalised measure on S^{n-1} . Since μ is known to be rationally invariant, the value of $d_\mu(y)$ is independent of $y \in S^{n-1}$. Hence μ is a d -invariant measure, and noting that the surface area of a sphere of radius r in \mathbb{R}^n is equal to $(2\pi^{n/2}r^{n-1})/\Gamma(n/2)$, it follows that

$$m(S^{n-1}, d) = \int_{S^{n-1}} d(x, e_n) d\mu(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{n/2}} \int_{S^{n-1}} \|x - e_n\| dx,$$

where $e_n = (0, \dots, 0, 1) \in \mathbb{R}^n$, and dx refers to integration with respect to the usual measure on S^{n-1} .

Let H be the upper hemisphere of S^{n-1} . As $\{x \in S^{n-1} : x_n = 0\}$ is a set of Lebesgue measure zero, we have that

$$\begin{aligned} \int_{S^{n-1}} \|x - e_n\| dx &= \int_H \|x - e_n\| dx + \int_H \|-x - e_n\| dx \\ &= \int_H \|x - e_n\| dx + \int_H \|x + e_n\| dx. \end{aligned}$$

Let U be the projection of H on \mathbb{R}^{n-1} and let $f: U \rightarrow \mathbb{R}$ be the function defined for $x' \in U$ by $f(x') = \sqrt{1 - \|x'\|^2}$. For each point $x = (x_1, \dots, x_n) \in H$, we may then write $x = (x', f(x'))$, where $x' = (x_1, \dots, x_{n-1}) \in U$. As ∇f exists for each point of U , it follows that

$$\begin{aligned} \int_H \|x - e_n\| dx &= \int_H \sqrt{\|x'\|^2 + (f(x') - 1)^2} dx \\ &= \int_U \left(\|x'\|^2 + \left(\sqrt{1 - \|x'\|^2} - 1 \right)^2 \right)^{1/2} \cdot \frac{1}{\sqrt{1 - \|x'\|^2}} dx' \\ &= \int_U \left(\frac{\|x'\|^2 + 1 - \|x'\|^2 - 2\sqrt{1 - \|x'\|^2} + 1}{1 - \|x'\|^2} \right)^{1/2} dx' \end{aligned}$$

$$= \int_U \left(\frac{2 - 2\sqrt{1 - \|x'\|^2}}{1 - \|x'\|^2} \right)^{1/2} dx',$$

where dx' refers to integration with respect to the usual measure on U .

For each $m \in \mathbb{N}$ and for each $r \geq 0$, let S_r^{m-1} denote the sphere of radius r centred at the origin in \mathbb{R}^m . Then the surface U is the union of all the spheres S_r^{n-2} such that $0 \leq r \leq 1$. By integrating over each S_r^{n-2} , we obtain

$$\begin{aligned} \int_U \left(\frac{2 - 2\sqrt{1 - \|x'\|^2}}{1 - \|x'\|^2} \right)^{1/2} dx' &= \int_0^1 \int_{S_r^{n-2}} \left(\frac{2 - 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} d\omega dr \\ &= \int_0^1 \left(\frac{2 - 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} \left(\int_{S_r^{n-2}} 1 d\omega \right) dr \\ &= \int_0^1 \left(\frac{2 - 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} \cdot \frac{2\pi^{(n-1)/2} r^{n-2}}{\Gamma\left(\frac{n-1}{2}\right)} dr \\ &= \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^1 r^{n-2} \left(\frac{2 - 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} dr. \end{aligned}$$

Similarly,

$$\begin{aligned} \int_H \|x + e_n\| dx &= \int_H \sqrt{\|x'\|^2 + (f(x') + 1)^2} dx \\ &= \int_U \left(\frac{2 + 2\sqrt{1 - \|x'\|^2}}{1 - \|x'\|^2} \right)^{1/2} dx' \\ &= \int_0^1 \int_{S_r^{n-2}} \left(\frac{2 + 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} d\omega dr \\ &= \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^1 r^{n-2} \left(\frac{2 + 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} dr. \end{aligned}$$

Using the substitution $r = \sin 2\theta$, where $0 \leq \theta \leq \pi/4$, it follows that

$$\begin{aligned} &\int_0^1 r^{n-2} \left(\left(\frac{2 - 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} + \left(\frac{2 + 2\sqrt{1 - r^2}}{1 - r^2} \right)^{1/2} \right) dr \\ &= \int_0^{\pi/4} \sin^{n-2} 2\theta \left(\left(\frac{2 - 2\cos 2\theta}{\cos^2 2\theta} \right)^{1/2} + \left(\frac{2 + 2\cos 2\theta}{\cos^2 2\theta} \right)^{1/2} \right) 2\cos 2\theta d\theta \end{aligned}$$

$$\begin{aligned}
&= 2 \int_0^{\pi/4} \sin^{n-2} \theta \left(\sqrt{4 \sin^2 \theta} + \sqrt{4 \cos^2 \theta} \right) d\theta \\
&= 2^n \int_0^{\pi/4} \sin^{n-2} \theta \cos^{n-2} \theta (\sin \theta + \cos \theta) d\theta.
\end{aligned}$$

Now, it is easily shown that the integrand of the previous integral is symmetric about $\pi/4$. Hence

$$\begin{aligned}
&2^n \int_0^{\pi/4} \sin^{n-2} \theta \cos^{n-2} \theta (\sin \theta + \cos \theta) d\theta \\
&= 2^{n-1} \int_0^{\pi/2} \sin^{n-2} \theta \cos^{n-2} \theta (\sin \theta + \cos \theta) d\theta \\
&= 2^{n-1} \int_0^{\pi/2} \sin^{n-1} \theta \cos^{n-2} \theta d\theta + 2^{n-1} \int_0^{\pi/2} \sin^{n-2} \theta \cos^{n-1} \theta d\theta,
\end{aligned}$$

and using the integrals

$$\int_0^{\pi/2} \sin^m x \cos^{m-1} x dx = \int_0^{\pi/2} \sin^{m-1} x \cos^m x dx = \frac{\Gamma\left(\frac{m+1}{2}\right) \Gamma\left(\frac{m}{2}\right)}{2\Gamma\left(\frac{2m+1}{2}\right)}$$

given by [44, pg. 402, 792] for $m \in \mathbb{N}$, it follows that

$$\int_0^1 r^{n-2} \left(\left(\frac{2 - 2\sqrt{1-r^2}}{1-r^2} \right)^{1/2} + \left(\frac{2 - 2\sqrt{1-r^2}}{1-r^2} \right)^{1/2} \right) dr = \frac{2^{n-1} \Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{2n-1}{2}\right)}.$$

We may now calculate

$$m(S^{n-1}, d) = \frac{\Gamma\left(\frac{n}{2}\right)}{2\pi^{n/2}} \cdot \frac{2\pi^{(n-1)/2}}{\Gamma\left(\frac{n-1}{2}\right)} \cdot \frac{2^{n-1} \Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{2n-1}{2}\right)} = \frac{2^{n-1} \left(\Gamma\left(\frac{n}{2}\right)\right)^2}{\sqrt{\pi} \Gamma\left(\frac{2n-1}{2}\right)}. \quad \square$$

The calculation of $m(S^{n-1})$ is based on the existence of the Lebesgue surface measure, which is known to be rotationally invariant and hence d -invariant on $m(S^{n-1})$, where d is the Euclidean metric. The next result demonstrates a construction of d -invariant measures on certain compact connected subspaces of (S^1, d) . The proof is found in Cleary et al. [16, pg. 266-267], who attributes the calculations to Des Robbie.

Proposition 4.5.8. *Let X_ϕ be an arc of S^1 subtending an angle $0 < \phi \leq 2\pi$ at the centre and let d be the usual metric on X_ϕ . Then there exists a d -invariant measure on X_ϕ and*

$$m(X_\phi) = \frac{8}{\phi + \cot(\phi/4)}.$$

Proof. Note that X_ϕ is a compact connected metric space. For all $0 \leq \theta \leq \phi$, let $P_\theta = (\cos \theta, \sin \theta)$. Then X_ϕ may be represented as $\{P_\theta : 0 \leq \theta \leq \phi\}$. Let $m \geq 0$, let $\mu \in \mathcal{M}^1(X_\phi)$ denote the usual normalised measure on X_ϕ and let μ_0 be the measure $\mu_0 = \mu + m\delta_{P_0} + m\delta_{P_\phi}$. We want to select m such that μ_0 is a d -invariant measure. (We will later normalise μ_0 to be a probability measure.)

Recall from Proposition 4.5.6 that for all $0 \leq \theta \leq \phi$, $\|P_\theta - P_0\| = 2 \sin(\theta/2)$. It follows by using a suitable rotation that for all $0 \leq \theta_1 \leq \theta_2 \leq \phi$,

$$\|P_{\theta_1} - P_{\theta_2}\| = 2 \sin \frac{\theta_2 - \theta_1}{2}.$$

Hence for all $0 \leq \alpha \leq \phi$,

$$\begin{aligned} & \int \|P - P_\alpha\| d\mu_0(P) \\ &= \int_0^\phi \|P_\theta - P_\alpha\| d\left(\frac{1}{2}\theta\right) + m\|P_0 - P_\alpha\| + m\|P_\phi - P_\alpha\| \\ &= \int_0^\alpha 2 \sin \frac{\alpha - \theta}{2} d\left(\frac{1}{2}\theta\right) + \int_\alpha^\phi 2 \sin \frac{\theta - \alpha}{2} d\left(\frac{1}{2}\theta\right) + 2m \sin \frac{\alpha}{2} + 2m \sin \frac{\phi - \alpha}{2} \\ &= 2 \left(1 - \cos \frac{\alpha}{2} - \cos \frac{\phi - \alpha}{2} + 1 + m \sin \frac{\alpha}{2} + m \sin \frac{\phi - \alpha}{2} \right) \\ &= 2 \left(2 + \left(m \sin \frac{\phi}{2} - \cos \frac{\phi}{2} - 1 \right) \cos \frac{\alpha}{2} + \left(m - m \cos \frac{\phi}{2} - \sin \frac{\phi}{2} \right) \sin \frac{\alpha}{2} \right). \end{aligned}$$

Therefore μ_0 will be a d -invariant measure if and only if

$$m \sin \frac{\phi}{2} - \cos \frac{\phi}{2} - 1 = m - m \cos \frac{\phi}{2} - \sin \frac{\phi}{2} = 0.$$

As

$$\frac{\sin x}{1 - \cos x} = \frac{\sin x}{1 - \cos x} \cdot \frac{1 + \cos x}{1 + \cos x} = \frac{\sin x(1 + \cos x)}{\sin^2 x} = \frac{1 + \cos x}{\sin x}$$

for all $x \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$, this condition is equivalent to

$$m = \frac{\sin(\phi/2)}{1 - \cos(\phi/2)}.$$

Note that for $0 < \phi \leq 2\pi$, the expression on the right-hand side of the above equality is defined and non-negative. Let m be chosen now as above. Then μ_0 is a d -invariant measure and $\mu_0(X_\phi) = \phi/2 + 2m$. Hence $\mu_0/(\phi/2 + 2m) \in \mathcal{M}^1(X_\phi)$ and we may calculate

$$\begin{aligned} m(X_\phi) &= \int \|P - P_0\| d\left(\frac{1}{\phi/2 + 2m}\mu_0\right)(P) \\ &= \frac{4}{\phi/2 + 2m} \\ &= \frac{8}{\phi + 4\sin(\phi/2)/(1 - \cos(\phi/2))}. \end{aligned}$$

The result follows from the identity

$$\frac{\sin(\phi/2)}{1 - \cos(\phi/2)} = \frac{2\sin(\phi/4)\cos(\phi/4)}{2 - 2\cos^2(\phi/4)} = \frac{\sin(\phi/4)\cos(\phi/4)}{\sin^2(\phi/4)} = \cot(\phi/4). \quad \square$$

This result leads to a surprising property of the Gross-Stadje number.

Theorem 4.5.9. *The mapping $m: \overline{H}(\mathbb{R}^2) \rightarrow \mathbb{R}$ is not monotonic with respect to subset inclusion.*

Proof. Let X_ϕ be the compact connected subspaces of S^1 , as defined in Proposition 4.5.8. Then for all $0 \leq \phi \leq 2\pi$,

$$\left. \frac{dm}{d\phi} \right|_\phi = \frac{d}{d\phi} \left(\frac{8}{\phi + \cot(\phi/4)} \right) = \frac{2\cot(\phi/4)^2 - 6}{(\phi + \cot(\phi/4))^2},$$

which is negative for $(2/3)\pi \leq \phi \leq 2\pi$. Let $\phi_1 = (2/3)\pi$ and let $\phi_2 = 2\pi$. Hence $X_{\phi_1} \subseteq X_{\phi_2}$ and $m(X_{\phi_1}) > m(X_{\phi_2})$. \square

We conclude this section by noting a result of Morris and Nickolas [38, pg. 461].

Proposition 4.5.10. *Let (G, d) be a compact connected metric group. Then*

$$m(G, d) = \int d(x, y) d\mu_0(x),$$

where $\mu_0 \in \mathcal{M}^1(G)$ denotes normalised Haar measure on G .

4.5.3 The support of max-min and M -maximal measures

We will make use of Caratheodory's Theorem, a proof of which is given by Eggleston [21, pg. 35–36].

Theorem 4.5.11. *Let $n \in \mathbb{N}$ and let $A \subseteq \mathbb{R}^n$. For each point x in the convex hull of A there exists $A_x \subseteq A$ such that $|A_x| \leq n + 1$ and x is in the convex hull of A_x .*

The following result is due to Peter Nickolas.

Theorem 4.5.12. *Let X be a compact subset of a Euclidean space. Then for all $\mu \in \mathcal{M}^1(X)$ there exists $\nu \in \mathcal{M}^1(X)$ concentrated on the closure of the extreme points of the convex hull of X such that $d_\mu \leq d_\nu$.*

Proof. Let $\mu \in \mathcal{M}^1(X)$. If μ is concentrated on the closure of the extreme points of the convex hull of X then there is nothing more to show. Otherwise, suppose that μ is an atomic measure. Let $x_0 \in X$ be an atom of μ which is not an extreme point of the convex hull of X and let E be the set of extreme points of the convex hull of X . Then by Caratheodory's Theorem there exist extreme points $x_1, \dots, x_n \in E$ and $\lambda_1, \dots, \lambda_n > 0$ such that $x_0 = \sum_{i=1}^n \lambda_i x_i$ and $\sum_{i=1}^n \lambda_i = 1$. Let $\nu_1, \nu_2 \in \mathcal{M}(X)$ be the positive measures defined for $A \in \mathcal{B}(X)$ by

$$\nu_1(A) = \mu(A \setminus \{x_0\}) \quad \text{and} \quad \nu_2(A) = \mu(\{x_0\}) \sum_{i=1}^n \lambda_i \delta_{x_i}(A),$$

and let $\nu = \nu_1 + \nu_2 \in \mathcal{M}^1(X)$. The measure ν spreads the mass given to x_0 by μ to the extreme points x_i with weighting given by the barycentric coordinates λ_i . Then for all $y \in X$,

$$\begin{aligned} d_\mu(y) &= \int_{X \setminus \{x_0\}} \|x - y\| d\mu(x) + \int_{\{x_0\}} \|x - y\| d\mu(x) \\ &= \int_{X \setminus \{x_0\}} \|x - y\| d\mu(x) + \mu(\{x_0\}) \|x_0 - y\| \\ &= \int_{X \setminus \{x_0\}} \|x - y\| d\mu(x) + \mu(\{x_0\}) \left\| \sum_{i=1}^n \lambda_i x_i - \sum_{i=1}^n \lambda_i y \right\| \end{aligned}$$

$$\begin{aligned}
&\leq \int_{X \setminus \{x_0\}} \|x - y\| d\mu(x) + \mu(\{x_0\}) \sum_{i=1}^n \lambda_i \|x_i - y\| \\
&= \int \|x - y\| d\nu_1(x) + \int \|x - y\| d\nu_2(x) \\
&= d_\nu(y).
\end{aligned}$$

The measure ν has one fewer non-extreme support point than μ . By repeating this process a finite number of times, there exists $\nu \in \mathcal{M}^1(X)$ such that ν is supported on the extreme points of X and $d_\mu \leq d_\nu$.

Suppose now that μ is an arbitrary probability measure. Let $(\mu_n) \in \mathcal{M}^1(X)$ be a convergent sequence of atomic measures such that $\mu_n \rightarrow^* \mu$, and let $(\nu_n) \in \mathcal{M}^1(X)$ be a sequence of atomic measures supported on E such that $d_{\mu_n} \leq d_{\nu_n}$ for each n . By the compactness of $\mathcal{M}^1(X)$ there exists $\nu \in \mathcal{M}^1(X)$ and a subsequence (ν_{n_k}) of (ν_n) such that $\nu_{n_k} \rightarrow^* \nu$. It follows that $\mu_{n_k} \rightarrow^* \mu$ and $d_\mu \leq d_\nu$. Let \overline{E} denote the closure of E . Then as for all k , $\nu_{n_k}(X \setminus \overline{E}) = 0$, it follows from Theorem 2.3.13 that

$$\nu(X \setminus \overline{E}) \leq \limsup_k \nu_{n_k}(X \setminus \overline{E}) = \limsup_k 0 = 0.$$

Hence, the support of ν is contained in \overline{E} . □

Corollary 4.5.13. *Let X be a compact connected subset of a Euclidean space. Then there exists a max-min Borel probability measure concentrated on the closure of the extreme points of the convex hull of X .*

Proof. Let $\mu, \nu \in \mathcal{M}^1(X)$ such that μ is an max-min measure and ν is supported on the closure of the extreme points of the convex hull of X and $d_\mu \leq d_\nu$. Then

$$\min_{y \in X} d_\mu(y) \leq m(X) \quad \text{and} \quad m(X) = \min_{y \in X} d_\mu(y) \leq \min_{y \in X} d_\nu(y). \quad \square$$

Corollary 4.5.14. *Let X be a compact subset of a Euclidean space. Then there exists an M -maximal Borel probability measure concentrated on the closure of the extreme points of the convex hull of X .*

Proof. Let $\mu, \nu \in \mathcal{M}^1(X)$ such that μ is an M -maximal measure and ν is supported on the closure of the extreme points of the convex hull of X and $d_\mu \leq d_\nu$. Then

$I(\nu) \leq M(X)$ and

$$M(X) = I(\mu) = \int d_\mu d\mu \leq \int d_\nu d\mu = \int d_\mu d\nu \leq \int d_\nu d\nu = I(\nu). \quad \square$$

We may use the previous result to calculate $M(X)$, where X is the perimeter of an equilateral triangle. In Chapter 6, we will use a similar technique to calculate an M -maximal measure on X , where X is the perimeter of a regular n -gon.

Proposition 4.5.15. *Let X be the perimeter of an equilateral triangle. Then the probability measure uniformly distributed on the vertices of X is an M -maximal measure and $M(X) = 2/3$.*

Proof. Let x_1, x_2, x_3 be the vertices of X . It follows from Corollary 4.5.14 that there exists an M -maximal measure $\mu \in \mathcal{M}^1(X)$ supported on the x_i . Hence, there exists $w_1, w_2, w_3 \geq 0$ such that $\mu = \sum_{i=1}^3 w_i \delta_{x_i}$ and $\sum_{i=1}^3 w_i = 1$. We have that

$$\begin{aligned} M(X) = I(\mu) &= \sum_{i=1}^3 \sum_{j=1}^3 w_i w_j \|x_i - x_j\| \\ &= \sum_{i=1}^3 \sum_{j=1}^3 w_i w_j \delta_{i,j} \\ &= 2w_1 w_2 + 2w_1 w_3 + 2w_2 w_3, \end{aligned}$$

where $\delta_{i,j}$ is the Kronecker delta. It will be sufficient to show that

$$\max(2w_1 w_2 + 2w_1 w_3 + 2w_2 w_3) = \frac{2}{3},$$

where the maximum is taken over all $w_1, w_2, w_3 \geq 0$ and $\sum_{i=1}^3 w_i = 1$.

Using the constraint $\sum_{i=1}^3 w_i = 1$, we have that

$$\begin{aligned} &2w_1 w_2 + 2w_1 w_3 + 2w_2 w_3 \\ &= 2w_1 w_2 + 2w_1(1 - w_1 - w_2) + 2w_2(1 - w_1 - w_2) \\ &= -2(w_1)^2 + (2 - 2w_2)w_1 + 2w_2(1 - w_2), \end{aligned}$$

which may be considered as a quadratic in x_1 whose value is maximised for a fixed x_2 when

$$x_1 = \frac{-(2 - 2w_2)}{-4} = \frac{1 - w_2}{2}.$$

Similarly, $2w_1w_2 + 2w_1w_3 + 2w_2w_3$ is maximised for a fixed x_1 by

$$x_2 = \frac{-(2 - 2w_1)}{-4} = \frac{1 - w_1}{2}.$$

To maximise this expression subject to our constraint, it is necessary that

$$x_1 = \frac{1 - w_2}{2} \quad \text{and} \quad x_2 = \frac{1 - w_1}{2} \quad \text{and} \quad w_1 + w_2 + w_3 = 1.$$

This system of equations has the unique solution $w_1 = w_2 = w_3 = 1/3$, giving that

$$M(X) = I(\mu) = 3 \times 2 \times \frac{1}{3} \times \frac{1}{3} = \frac{2}{3}. \quad \square$$

4.6 Summary

Chapters 2 and 3 discussed measure theory which was used in this chapter to give additional characterisations of $m(X)$, $M(X)$ and $\overline{M}(X)$, and to develop a greater understanding of these average distances. In our discussion of $m(X)$, we supplied a measure theoretic proof of the Gross-Stadje Theorem. Such an argument was omitted from the original discussion of this theorem in Chapter 1.

Our measure theoretic characterisations of $m(X)$, $M(X)$ and $\overline{M}(X)$ were found to be useful for investigating properties of these average distances. For example, using max-min measures it is clear that $m(X) \leq M(X)$, a result which is otherwise harder to establish. Other properties which were discussed involved the range of permissible values for $m(X)$, $M(X)$ and $\overline{M}(X)$, the existence of max-min and min-max measures which give rise to the value of $m(X)$, the existence of M -maximal measures which give rise to the value of $M(X)$, and properties of $m(X)$, $M(X)$ and $\overline{M}(X)$ when considered as real-valued functions. Little was said in general of $\overline{M}(X)$, as we will discover in Chapter 6 that this average distance is easier understood when considered along with the so-called quasihypermetric inequality.

We concluded with a demonstration of several techniques which may be used to calculate $m(X)$ and $M(X)$ for several compact (connected) metric spaces, but we did not calculate $\overline{M}(X)$ for any space. We will find some values of this later average distance in Chapter 6.

Chapter 5

Quasihypermetric spaces

This chapter introduces hypermetric, quasihypermetric and strictly quasihypermetric spaces. Our main results give concrete examples of each type of space, explore the relationship between them, and list some of their characterisations, paying particular attention to cases of compactness.

Although the development of the results in this chapter are motivated by their importance to our study of average distances, it is meant to be a self-contained unit which may be read independently of the preceding work in this thesis. We shall see in Chapter 6 that considering quasihypermetric spaces gives us a greater understanding of average distances in compact metric spaces.

5.1 Hypermetric spaces

Definition 5.1.1. Let (X, d) be a metric space and let $n \in \mathbb{N}$. If

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^{n+1} d(a_i, b_j)$$

for all $a_1, \dots, a_n, b_1, \dots, b_{n+1} \in X$, then (X, d) is an *n-hypermetric space*. If (X, d) is *n-hypermetric* for all $n \in \mathbb{N}$, then (X, d) is a *hypermetric space*.

The terms *n-hypermetric* and *hypermetric* were introduced by Kelly [28, pg. 18], who found these inequalities useful for studying the metric space of finite subsets of a

set, where the distance between two sets is calculated as the cardinality of their symmetric difference. According to Kelly, the first person to formulate the hypermetric inequality was Tylkin [51], who referred to it as the *f-polygonal inequality*.

We note that Kelly actually defines the n -hypermetric inequality to be

$$\sum_{1 \leq i < j \leq n} d(a_i, a_j) + \sum_{1 \leq i < j \leq n+1} d(b_i, b_j) \leq \sum_{1 \leq i \leq n, 1 \leq j \leq n+1} d(a_i, b_j),$$

and uses this inequality to classify semi-metric spaces. Given that this thesis only considers metric spaces, our definition of the n -hypermetric inequality is equivalent to Kelly's and is more useful for our purposes.

Noting that the 1-hypermetric inequality is equivalent to the triangle inequality, every metric space is a 1-hypermetric space.

5.1.1 Characterisations of hypermetric spaces

The following characterisations of hypermetric spaces concern weighted average distances. Characterisation 3 is stated in terms of atomic measures for ease of notation. The proof that (1) \iff (2) in Theorem 5.1.2 is suggested by Kelly [28, pg. 19]. The assertion that (2) \iff (3) is original.

Theorem 5.1.2. *Let (X, d) be a metric space. The following are equivalent:*

1. (X, d) is hypermetric.
2. For all $n \in \mathbb{N}$, for all $w_1, \dots, w_n \in \mathbb{Z}$ and for all $x_1, \dots, x_n \in X$,

$$\sum_{i=1}^n w_i = 1 \implies \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0.$$

3. For all $\mu \in \mathcal{M}(X)$, if $\mu(X) = 0$ and μ can be written as a finite linear combination of point measures using non-zero rational weights then

$$I(\mu) \leq -\frac{2}{m} \|d_\mu\|,$$

where m is the lowest common positive multiple of the denominators of the weights of μ .

Proof. To show that (1) \implies (2), suppose that (X, d) is hypermetric. Let $n \in \mathbb{N}$, let $w_1, \dots, w_n \in \mathbb{Z}$ such that $\sum_{i=1}^n w_i = 1$ and let $x_1, \dots, x_n \in X$. If each weight is non-negative then there exists k such that $w_k = 1$ and $w_i = 0$ for all $i \neq k$, and it follows that

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) = w_k w_k d(x_k, x_k) = 0.$$

Otherwise, there exists at least one negative weight and at least one positive weight. Let m be the number of negative weights, let $a_1, \dots, a_m \in X$ represent the x_i such that $w_i < 0$ and let $u_1, \dots, u_m \in \mathbb{Z}$ be the corresponding weights, and let $b_1, \dots, b_{n-m} \in X$ represent the x_i such that $w_i \geq 0$ and let $v_1, \dots, v_{n-m} \in \mathbb{Z}$ be the corresponding weights. Then

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \\ &= \sum_{i=1}^m \sum_{j=1}^m u_i u_j d(a_i, a_j) + \sum_{i=1}^{n-m} \sum_{j=1}^{n-m} v_i v_j d(b_i, b_j) + 2 \sum_{i=1}^m \sum_{j=1}^{n-m} u_i v_j d(a_i, b_j) \\ &= \sum_{i=1}^m \sum_{j=1}^m |u_i| \cdot |u_j| \cdot d(a_i, a_j) + \sum_{i=1}^{n-m} \sum_{j=1}^{n-m} |v_i| \cdot |v_j| \cdot d(b_i, b_j) \\ & \quad - 2 \sum_{i=1}^m \sum_{j=1}^{n-m} |u_i| \cdot |v_j| \cdot d(a_i, b_j). \end{aligned}$$

Let $m' = \sum_{i=1}^m -u_i$ and let $n' = \sum_{i=1}^{n-m} v_i$. Since $-m' + n' = 1$, it follows that $n' = m' + 1$. Let $a'_1, \dots, a'_{m'} \in X$ be the list of points consisting of exactly $|u_i|$ occurrences of a_i for each i and let $b'_1, \dots, b'_{m'+1} \in X$ be the list of points consisting of exactly $|v_i|$ occurrences of b_i for each i . Then as (X, d) is hypermetric,

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \\ &= \sum_{i=1}^{m'} \sum_{j=1}^{m'} d(a'_i, a'_j) + \sum_{i=1}^{m'+1} \sum_{j=1}^{m'+1} d(b'_i, b'_j) - 2 \sum_{i=1}^{m'} \sum_{j=1}^{m'+1} d(a'_i, b'_j) \\ &\leq 0. \end{aligned}$$

To show that (2) \implies (1), suppose that for all $n \in \mathbb{N}$, for all $w_1, \dots, w_n \in \mathbb{Z}$ and

for all $x_1, \dots, x_n \in X$,

$$\sum_{i=1}^n w_i = 1 \implies \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0.$$

Let $n \in \mathbb{N}$, let $a_1, \dots, a_n, b_1, \dots, b_{n+1} \in X$, let $x_1, \dots, x_{2n+1} \in X$ be a concatenated listing of the points a_i and b_i , and let $w_1, \dots, w_{2n+1} \in \mathbb{Z}$ such that $w_i = -1$ if $i \leq n$ and $w_i = 1$ if $i > n$. We have that $\sum_{i=1}^{2n+1} w_i = 1$, giving

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} d(b_i, b_j) - 2 \sum_{i=1}^n \sum_{j=1}^{n+1} d(a_i, b_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n (-1) \cdot (-1) \cdot d(a_i, a_j) + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} 1 \cdot 1 \cdot d(b_i, b_j) \\ & \quad + 2 \sum_{i=1}^n \sum_{j=1}^{n+1} (-1) \cdot 1 \cdot d(a_i, b_j) \\ &= \sum_{i=1}^{2n+1} \sum_{j=1}^{2n+1} w_i w_j d(x_i, x_j) \\ &\leq 0. \end{aligned}$$

Hence

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^{n+1} d(a_i, b_j),$$

giving that (X, d) is hypermetric.

To show that (2) \implies (3), suppose that $I(\nu) \leq 0$ for all $\nu \in \mathcal{M}(X)$ such that $\nu(X) = 1$ and ν can be written as a finite linear combination of point measures using integer weights. (This is condition (2), stated in our measure theory terminology.) Let $\mu \in \mathcal{M}(X)$ be such that $\mu(X) = 0$ and μ can be written as a finite linear combination of point measures using non-zero rational weights, let $m \neq 0$ be the lowest common positive multiple of the denominators of the weights of μ and let $x \in X$. Since the signed measure $m\mu + \delta_x$ is a finite linear combination of point measures using integer weights and $(m\mu + \delta_x)(X) = 1$, it follows that $I(m\mu + \delta_x) \leq 0$, which then gives $m^2 I(\mu) + I(\delta_x) + 2mI(\mu, \delta_x) \leq 0$ and

$$I(\mu) \leq -\frac{2}{m} d_\mu(x).$$

By applying the above argument to $-\mu$, given that $d_{-\mu} = -d_\mu$ and $I(-\mu) = I(\mu)$ we obtain

$$I(\mu) \leq -\frac{2}{m}d_{-\mu}(x) = \frac{2}{m}d_\mu(x),$$

from which it follows that $I(\mu) \leq -(2/m)|d_\mu(x)|$. As this is true for all $x \in X$, we then have

$$I(\mu) \leq -\frac{2}{m}\|d_\mu\|.$$

To show that (3) \implies (2), suppose that for all $\nu \in \mathcal{M}(X)$, if $\nu(X) = 0$ and ν can be written as a finite linear combination of point measures using non-zero rational weights then

$$I(\nu) \leq -\frac{2}{m}\|d_\nu\|,$$

where m is the lowest common positive multiple of the denominators of the weights of ν . Let $\mu \in \mathcal{M}(X)$ be such that $\mu(X) = 1$ and μ can be written as a finite linear combination of point measures using integer weights, and let $x \in X$. We then have $(\mu - \delta_x)(X) = 0$ and by taking $m = 1$, since $I(\mu - \delta_x) \leq -2\|d_{\mu - \delta_x}\| \leq -2d_{\mu - \delta_x}(x)$, it follows that $I(\mu) + I(\delta_x) - 2I(\mu, \delta_x) \leq -2d_\mu(x)$ and hence $I(\mu) \leq 0$, which is the conclusion of (2). \square

We then obtain:

Corollary 5.1.3. *Let (X, d) be a hypermetric space. Then for all $\mu \in \mathcal{M}(X)$, if $\mu(X) = 0$ and μ can be written as a finite linear combination of point measures using non-zero rational weights then $I(\mu) = 0 \iff d_\mu = 0$.*

Proof. Let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$ and μ can be written as a finite linear combination of point measures using rational weights. Then

$$d_\mu = 0 \implies \int d_\mu d\mu = 0 \implies I(\mu) = 0$$

and

$$d_\mu \neq 0 \implies \|d_\mu\| > 0 \implies I(\mu) \leq -\frac{2}{m}\|d_\mu\| < 0,$$

where m is some positive constant dependent on μ . \square

5.1.2 Examples of hypermetric spaces

We firstly show that \mathbb{R}^n is hypermetric for all $n \in \mathbb{N}$. Our proof that \mathbb{R} is hypermetric uses essentially the same argument as that which proved Theorem 1.2.7, though we shall see in Section 5.2 that the property of \mathbb{R} proved in this referenced theorem is a weaker characterisation of \mathbb{R} than the hypermetric inequality.

The proof of the following result is due to Kézdy et al. [30, pg. 25-26].

Theorem 5.1.4. *\mathbb{R} is hypermetric.*

Proof. Let $n \in \mathbb{N}$ and as in the proof of Theorem 1.2.7, define the constants α_k, β_k and γ_k , and t_k, u_k, y_k and z_k , the only difference being that the list of points b_i which lead to the definition of these constants should be of length $n+1$. Then $t_k + y_k = n$ and $u_k + z_k = n+1$ and it can be shown that $\alpha_k = t_k y_k$ and $\beta_k = u_k z_k$ and $\gamma_k = t_k z_k + u_k y_k$. It follows from $(u_k - t_k) - (y_k - z_k) = (u_k + z_k) - (t_k + y_k) = n+1 - n = 1$ that $u_k - t_k = (y_k - z_k) + 1$, and given that $c(c+1) \geq 0$ for all $c \in \mathbb{Z}$, we then have $(u_k - t_k)(y_k - z_k) \geq 0$. Hence $t_k y_k + u_k z_k \leq t_k z_k + u_k y_k$, giving $\alpha_k + \beta_k \leq \gamma_k$. By the definition of α_k, β_k and γ_k , it follows that the n -hypermetric inequality holds in \mathbb{R} , and hence \mathbb{R} is hypermetric. \square

Kelly [28, pg. 29-30] investigated under what conditions equality holds in the hypermetric inequality for \mathbb{R} . We state his result without proof.

Proposition 5.1.5. *Let $n, m \in \mathbb{N}$ and let $a_1, \dots, a_m, b_1, \dots, b_{m+1} \in \mathbb{R}^n$. Then equality will occur in the hypermetric inequality for the a_i and b_i if and only if the a_i interleave the b_i ; that is, by assuming the a_i and b_i to be ordered,*

$$b_1 \leq a_1 \leq b_2 \leq a_2 \leq b_m \leq a_m \leq b_{m+1}.$$

An outline proof of the following result is found in Kelly [26, pg. 201].

Lemma 5.1.6. *Let $n \in \mathbb{N}$ such that $n \geq 2$ and let $y \in \mathbb{R}^n$. Then*

$$\int_{S^{n-1}} |\langle x, y \rangle| dx = 2 \cdot \|y\| \cdot V_{n-1},$$

where dx refers to integration with respect to the usual surface measure on S^{n-1} and V_{n-1} is the volume of the unit ball in \mathbb{R}^{n-1} .

Proof. If $y = 0$ the result is obvious. Otherwise, $y \neq 0$ and

$$\int_{S^{n-1}} |\langle x, y \rangle| dx = \|y\| \int_{S^{n-1}} \left| \left\langle x, \frac{y}{\|y\|} \right\rangle \right| dx.$$

By the symmetry of S^{n-1} , the value of the integral on the right-hand side of the above equation is independent of $y/\|y\| \in S^{n-1}$, and by writing each $x \in S^{n-1}$ as $x = (x_1, \dots, x_n)$ it follows that

$$\int_{S^{n-1}} |\langle x, y \rangle| dx = \|y\| \int_{S^{n-1}} |\langle x, e_n \rangle| dx = \|y\| \int_{S^{n-1}} |x_n| dx,$$

where $e_n = (0, \dots, 0, 1) \in S^{n-1}$.

Let H denote the upper hemisphere of S^{n-1} . As $\{x \in S^{n-1} : x_n = 0\}$ is a set of Lebesgue measure zero, we have that

$$\int_{S^{n-1}} |x_n| dx = \int_H |x_n| dx + \int_H |-x_n| dx = 2 \int_H x_n dx.$$

Let U be the projection of H on \mathbb{R}^{n-1} , which is of course the unit ball in \mathbb{R}^{n-1} , and let $f: U \rightarrow \mathbb{R}$ be the function defined for $x' \in U$ by $f(x') = \sqrt{1 - \|x'\|^2}$. For each $x \in H$, we may then write $x = (x', f(x'))$, where $x' = (x_1, \dots, x_{n-1}) \in U$. As ∇f exists for each point of U , it follows that

$$\int_H x_n dx = \int_U \sqrt{1 - \|x'\|^2} \cdot \frac{1}{\sqrt{1 - \|x'\|^2}} dx' = \int_U 1 dx' = V_{n-1},$$

where dx' refers to integration with respect to the Lebesgue measure on U . Hence

$$\int_{S^{n-1}} |\langle x, y \rangle| dx = 2 \cdot \|y\| \cdot V_{n-1}. \quad \square$$

The following proof is due to Kézdy et al. [30, pg. 26-27].

Theorem 5.1.7. *Let $n \in \mathbb{N}$. Then \mathbb{R}^n is hypermetric.*

Proof. Theorem 5.1.4 gives the result for $n = 1$, so we will assume that $n \geq 2$. Let $m \in \mathbb{N}$, let $a_1, \dots, a_m, b_1, \dots, b_{m+1} \in \mathbb{R}^n$ and let $f: S^{n-1} \rightarrow \mathbb{R}$ and $g: S^{n-1} \rightarrow \mathbb{R}$ be the functions defined for $x \in S^{n-1}$ by

$$\begin{aligned} f(x) &= \sum_{i=1}^m \sum_{j=1}^m |\langle x, a_i \rangle - \langle x, a_j \rangle| + \sum_{i=1}^{m+1} \sum_{j=1}^{m+1} |\langle x, b_i \rangle - \langle x, b_j \rangle|, \\ g(x) &= 2 \sum_{i=1}^m \sum_{j=1}^{m+1} |\langle x, a_i \rangle - \langle x, b_j \rangle|. \end{aligned}$$

Noting that $\langle x, a_i \rangle, \langle x, b_j \rangle \in \mathbb{R}$ for all $x \in S^{n-1}$ and all i and j , Theorem 5.1.4 gives that $f \leq g$. Now for all $y \in \mathbb{R}$, by Lemma 5.1.6 we have

$$\|y\| = \frac{1}{2V_{n-1}} \int_{S^{n-1}} |\langle x, y \rangle| dx,$$

where V_{n-1} is the volume of the unit ball in \mathbb{R}^{n-1} . Therefore

$$\begin{aligned} & \sum_{i=1}^m \sum_{j=1}^m \|a_i - a_j\| + \sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \|b_i - b_j\| \\ &= \sum_{i=1}^m \sum_{j=1}^m \frac{1}{2V_{n-1}} \int_{S^{n-1}} |\langle x, a_i \rangle - \langle x, a_j \rangle| dx \\ & \quad + \sum_{i=1}^{m+1} \sum_{j=1}^{m+1} \frac{1}{2V_{n-1}} \int_{S^{n-1}} |\langle x, b_i \rangle - \langle x, b_j \rangle| dx \\ &= \frac{1}{2V_{n-1}} \int_{S^{n-1}} \left(\sum_{i=1}^m \sum_{j=1}^m |\langle x, a_i \rangle - \langle x, a_j \rangle| + \sum_{i=1}^{m+1} \sum_{j=1}^{m+1} |\langle x, b_i \rangle - \langle x, b_j \rangle| \right) dx \\ &= \frac{1}{2V_{n-1}} \int_{S^{n-1}} f(x) dx \\ &\leq \frac{1}{2V_{n-1}} \int_{S^{n-1}} g(x) dx \\ &= 2 \sum_{i=1}^m \sum_{j=1}^{m+1} \|a_i - b_j\|. \end{aligned}$$

Hence \mathbb{R}^n is hypermetric. □

The fact that \mathbb{R}^n is hypermetric can also be used to show that several of the L^p spaces are hypermetric. Recall that given a measure space (X, \mathcal{X}, μ) and $p \geq 1$, $L^p = L^p(X, \mathcal{X}, \mu)$ consists of all integrable functions $f: X \rightarrow \mathbb{R}$ such that $|f|^p$ is integrable, and is equipped with the semi-norm $\|\cdot\|_p: L^p \rightarrow \mathbb{R}$ defined for $f \in L^p$ by

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}.$$

By defining an equivalence relation on L^p relating functions which are equal μ -almost everywhere, it is customary to consider L^p to be a normed linear space, and with this convention it can be shown that L^p is complete. Further, when $p = 2$, L^2 is a Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle: L^2 \rightarrow \mathbb{R}$ defined for $f, g \in L^2$ by

$$\langle f, g \rangle = \int fg d\mu.$$

Kelly [28, pg.28] notes that L^1 and L^2 spaces are hypermetric. We now derive these results as a consequence of Theorem 5.1.4.

In the following proofs and for the remainder of this thesis, we find it convenient to let \mathbb{N}_0 denote $\mathbb{N} \cup \{0\}$, to let \mathbb{R}^0 denote the trivial metric space $\{0\}$, to let \mathbb{R}^1 denote the Euclidean space \mathbb{R} , and to let $\text{sp } X$ denote the vector space spanned by some subset X of a vector space.

The following result is well-known.

Lemma 5.1.8. *Let V be an inner product space, let $n \in \mathbb{N}_0$ and let $v_0, \dots, v_n \in V$. Then the v_i can be isometrically embedded in \mathbb{R}^m for some $m \leq n$, where m is the greatest number of independent points amongst the v_i .*

Proof. By considering a suitable translation, we may assume without loss of generality that $v_0 = 0$. Let $W = \text{sp } \{v_1, \dots, v_n\}$. Then W is a finite dimensional subspace of V of dimension $m \leq n$, which is the greatest number of independent points amongst the v_i . By using the Gram-Schmidt process let d_1, \dots, d_m be an orthonormal basis for W , let e_1, \dots, e_m be the standard basis for \mathbb{R}^m and let $T: W \rightarrow \mathbb{R}^m$ be the linear transformation defined by $T(d_i) = e_i$ for each i . Then for each $x, y \in W$,

$$\begin{aligned}
 \|T(x) - T(y)\| &= \left\| T \left(\sum_{i=1}^m \langle x, d_i \rangle d_i \right) - T \left(\sum_{i=1}^m \langle y, d_i \rangle d_i \right) \right\| \\
 &= \left\| \sum_{i=1}^m \langle x, d_i \rangle e_i - \sum_{i=1}^m \langle y, d_i \rangle e_i \right\| \\
 &= \left\| \sum_{i=1}^m \langle x - y, d_i \rangle e_i \right\| \\
 &= \left(\sum_{i=1}^m \langle x - y, d_i \rangle^2 \right)^{1/2} \\
 &= \left\| \sum_{i=1}^m \langle x - y, d_i \rangle d_i \right\| \\
 &= \|x - y\|.
 \end{aligned}$$

Therefore T is an isometry, and the $T(v_i)$ are an embedding of the v_i in \mathbb{R}^m . \square

Theorem 5.1.9. *Let (X, \mathcal{X}, μ) be a measure space and let $p \in [1, 2]$. Then $L^p(X, \mathcal{X}, \mu)$ is hypermetric.*

Proof. Let $n \in \mathbb{N}$ and let $f_1, \dots, f_n, g_1, \dots, g_{n+1} \in L^1(X, \mathcal{X}, \mu)$. Then Theorem 5.1.4 gives that

$$\begin{aligned}
& \sum_{i=1}^n \sum_{j=1}^n \|f_i - f_j\|_1 + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \|g_i - g_j\|_1 \\
&= \sum_{i=1}^n \sum_{j=1}^n \int |f_i - f_j| d\mu + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \int |g_i - g_j| d\mu \\
&= \int \left(\sum_{i=1}^n \sum_{j=1}^n |f_i(x) - f_j(x)| + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} |g_i(x) - g_j(x)| \right) d\mu(x) \\
&\leq \int 2 \sum_{i=1}^n \sum_{j=1}^{n+1} |f_i(x) - g_j(x)| d\mu(x) \\
&= 2 \sum_{i=1}^n \sum_{j=1}^{n+1} \|f_i - g_j\|_1.
\end{aligned}$$

Hence $L^1(X, \mathcal{X}, \mu)$ is hypermetric. The result in the case that $p \in [1, 2]$ follows since then $L^p(X, \mathcal{X}, \mu)$ is isometric to a subspace of $L^1(X, \mathcal{X}, \mu)$, as noted by Koldobsky and Lonke [31, pg. 694], who refer to a paper in French by Bretagnolle et. al. [11]. It is easily seen directly from Lemma 5.1.8 that $L^2(X, \mathcal{X}, \mu)$ is hypermetric. \square

That $L^1(X, \mathcal{X}, \mu)$ is hypermetric leads to a test for hypermetric spaces given by Kelly [28, pg. 19-20]. Kelly proves this result by using an elementary argument, and then uses this test to demonstrate that \mathbb{R} is hypermetric [28, pg. 21]. We instead show that the result follows from Theorem 5.1.9, which is in turn a consequence of the hypermetric property for \mathbb{R} .

Proposition 5.1.10. *Let (X, d) be a metric space, let (Y, \mathcal{Y}, μ) be a measure space and let $\Delta: \mathcal{Y} \times \mathcal{Y} \rightarrow \mathcal{Y}$ be the symmetric difference operator. If there exists a function $\dot{f}: X \rightarrow \mathcal{Y}$ such that for all $x_1, x_2 \in X$, $d(x_1, x_2) = \mu(f(x_1)\Delta f(x_2))$, then (X, d) is hypermetric.*

Proof. Suppose that there exists a function $f : X \rightarrow \mathcal{Y}$ such that for all $x_1, x_2 \in X$, $d(x_1, x_2) = \mu(f(x_1)\Delta f(x_2))$. Since $|\chi_A - \chi_B| = \chi_{A\Delta B}$ for all $A, B \in \mathcal{Y}$, we have that for all $x_1, x_2 \in X$,

$$\begin{aligned} d(x_1, x_2) &= \mu(f(x_1)\Delta f(x_2)) \\ &= \int \chi_{f(x_1)\Delta f(x_2)} d\mu \\ &= \int |\chi_{f(x_1)} - \chi_{f(x_2)}| d\mu \\ &= \|\chi_{f(x_1)} - \chi_{f(x_2)}\|_1, \end{aligned}$$

which is the statement that the mapping $X \rightarrow L^1(Y, \mathcal{Y}, \mu)$ such that $x \mapsto \chi_{f(x)}$ is an isometry. It follows from Theorem 5.1.9 that (X, d) is hypermetric. \square

When $X = \mathbb{N}$, \mathcal{X} is the discrete Borel σ -algebra on \mathbb{N} , μ is counting measure on (X, \mathcal{X}) and $p = 2$, we obtain the well known Hilbert space of “square-summable sequences”

$$\ell_2 = L^2(\mathbb{N}, \mathcal{X}, \mu) = \left\{ (u_1, u_2, u_3, \dots) : u_i \in \mathbb{R} \text{ for each } i \text{ and } \sum_{i=1}^{\infty} u_i^2 \text{ converges} \right\},$$

which may be equipped with the inner product $\langle \cdot, \cdot \rangle : \ell_2 \times \ell_2 \rightarrow \mathbb{R}$ defined for $(u_1, u_2, u_3, \dots), (v_1, v_2, v_3, \dots) \in \ell_2$ by

$$\langle (u_1, u_2, u_3, \dots), (v_1, v_2, v_3, \dots) \rangle = \sum_{i=1}^{\infty} u_i v_i.$$

In our discussion of quasihypermetric spaces, we will see that ℓ_2 is an important hypermetric and quasihypermetric space. It is immediate that:

Theorem 5.1.11. *ℓ_2 is hypermetric.*

Using a result of Lindenstrauss, Kelly [28, pg.28] notes that all normed linear spaces of dimension no greater than 2 are hypermetric. It is easily seen that not every normed linear space is hypermetric. The proof of the following result is due to Kelly [26, pg. 202-203].

Proposition 5.1.12. *The normed linear space \mathbb{R}^3 equipped with the metric induced by the ∞ -norm is not 2-hypermetric, and consequently not hypermetric.*

Proof. Let $n = 2$ and let $a_1, a_2, b_1, b_2, b_3 \in \mathbb{R}^3$ be given by

$$\begin{aligned} a_1 &= (0, 0, 0), & a_2 &= (0, 0, 1), \\ b_1 &= (1, 1, 0), & b_2 &= (1, -1, 0), & b_3 &= (-1, 1, 0). \end{aligned}$$

Then

$$\begin{aligned} \|a_1 - a_2\|_\infty &= 1, \\ \|b_i - b_j\|_\infty &= 2 \quad \text{for all } i \text{ and all } j \text{ such that } i \neq j, \\ \|a_i - b_j\|_\infty &= 1 \quad \text{for all } i \text{ and } j, \end{aligned}$$

and

$$14 = \sum_{i=1}^n \sum_{j=1}^n \|a_i - a_j\|_\infty + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \|b_i - b_j\|_\infty > 2 \sum_{i=1}^n \sum_{j=1}^{n+1} \|a_i - b_j\|_\infty = 12. \quad \square$$

5.2 Quasihypermetric spaces

Definition 5.2.1. Let (X, d) be a metric space and let $n \in \mathbb{N}$. If

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j)$$

for all $a_1, \dots, a_n, b_1, \dots, b_n \in X$, then (X, d) is an n -quasihypermetric space. If (X, d) is n -quasihypermetric for all $n \in \mathbb{N}$, then (X, d) is a quasihypermetric space.

The term *quasihypermetric* was introduced by Kelly [28, pg. 26], but defined using a different characterisation of such spaces. We have chosen our definition of quasihypermetric spaces to be consistent with (our equivalent version of) Kelly's definition of hypermetric spaces. Kelly does not define the term n -quasihypermetric (as compared to the term n -hypermetric), but we have introduced it to help show the relationship between quasihypermetric and hypermetric spaces.

Noting that the 1-quasihypermetric inequality is equivalent to the statement that a metric is non-negative, every metric space is a 1-quasihypermetric space. Also, by considering the triangle inequality, a cases argument shows that every metric space is a 2-quasihypermetric space.

5.2.1 Relationships with hypermetric spaces

The results and comments in this section establish relationships between the class of hypermetric spaces and the class of quasihypermetric spaces.

Theorem 5.2.2. *Let (X, d) be a metric space. For each $n \in \mathbb{N}$, if (X, d) is n -hypermetric then (X, d) is n -quasihypermetric. Further, if (X, d) is hypermetric then (X, d) is quasihypermetric.*

Proof. Let $n \in \mathbb{N}$ and suppose that (X, d) is n -hypermetric. Let $a_1, \dots, a_n \in X$, let $b_1, \dots, b_n \in X$ and let $a_{n+1}, b_{n+1} \in X$ such that $a_{n+1} = b_{n+1}$. Since (X, d) is n -hypermetric, we have that

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^{n+1} d(a_i, b_j),$$

giving

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) + 2 \sum_{i=1}^n d(b_i, b_{n+1}) \\ & \leq 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j) + 2 \sum_{i=1}^n d(a_i, b_{n+1}). \end{aligned}$$

Similarly,

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) + 2 \sum_{i=1}^n d(a_i, a_{n+1}) \\ & \leq 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j) + 2 \sum_{i=1}^n d(b_i, a_{n+1}), \end{aligned}$$

and as

$$\sum_{i=1}^n d(b_i, b_{n+1}) = \sum_{i=1}^n d(b_i, a_{n+1}) \quad \text{and} \quad \sum_{i=1}^n d(a_i, b_{n+1}) = \sum_{i=1}^n d(a_i, a_{n+1}),$$

it follows that

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j).$$

Hence (X, d) is n -quasihypermetric. That (X, d) is quasihypermetric when (X, d) is hypermetric is immediate. \square

It is then immediate that:

Theorem 5.2.3. *Let $n \in \mathbb{N}_0$ and let (X, \mathcal{X}, μ) be a measure space. Then each of \mathbb{R}^n , $L^1(X, \mathcal{X}, \mu)$, $L^2(X, \mathcal{X}, \mu)$ and ℓ_2 are quasihypermetric.*

The following result is due to Witsenhausen [55, pg. 518]. We omit the proof.

Proposition 5.2.4. *All normed linear quasihypermetric spaces are hypermetric.*

It from Proposition 5.1.12 that:

Proposition 5.2.5. *The normed linear space \mathbb{R}^3 equipped with the metric induced by the ∞ -norm is not quasihypermetric.*

In Proposition 5.3.12 we will see an example of a five-point quasihypermetric space which is not hypermetric. Hence the class of quasihypermetric spaces is a proper subclass of the space of hypermetric spaces, justifying the use of the prefix “quasi” in the term “quasihypermetric”.

5.2.2 A characterisation of quasihypermetric spaces

Our main characterisation of quasihypermetric spaces uses weighted average distances. The proof of Proposition 5.2.6 is suggested by Lovász et al. [35, pg. 2049].

Theorem 5.2.6. *Let (X, d) be a metric space. Then (X, d) is quasihypermetric if and only if for all $n \in \mathbb{N}$, for all $w_1, \dots, w_n \in \mathbb{R}$ and for all $x_1, \dots, x_n \in X$,*

$$\sum_{i=1}^n w_i = 0 \implies \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0.$$

Proof. Suppose that

$$\sum_{i=1}^n w_i = 0 \implies \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0$$

for all $n \in \mathbb{N}$, for all $w_1, \dots, w_n \in \mathbb{R}$ and for all $x_1, \dots, x_n \in X$. Let $n \in \mathbb{N}$, let $a_1, \dots, a_n, b_1, \dots, b_n \in X$, let $x_1, \dots, x_{2n} \in X$ be a listing of the points a_i and b_i , and let $w_1, \dots, w_{2n} \in \mathbb{Z}$ such that $w_i = -1$ if $i \leq n$ and $w_i = 1$ if $i > n$. We have that $\sum_{i=1}^{2n} w_i = 0$, and therefore

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) - 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n (-1)(-1)d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n (1)(1)d(b_i, b_j) + 2 \sum_{i=1}^n \sum_{j=1}^n (-1)(1)d(a_i, b_j) \\ &= \sum_{i=1}^{2n} \sum_{j=1}^{2n} w_i w_j d(x_i, x_j) \\ &\leq 0. \end{aligned}$$

Since then

$$\sum_{i=1}^n \sum_{j=1}^n d(a_i, a_j) + \sum_{i=1}^n \sum_{j=1}^n d(b_i, b_j) \leq 2 \sum_{i=1}^n \sum_{j=1}^n d(a_i, b_j),$$

it follows that (X, d) is quasihypermetric. The proof of the converse statement is identical to that of Theorem 1.2.9. \square

5.3 Compact quasihypermetric spaces

We now develop characterisations of compact quasihypermetric spaces which fall into two categories. The first collection of characterisations are measure theoretic, and the second collection of characterisations concern certain embeddings which are essentially in ℓ_2 .

5.3.1 Measure theoretic characterisations

Theorem 5.3.1. *Let X be a compact metric space. The following are equivalent:*

1. X is quasihypermetric.
2. For all $\mu \in \mathcal{M}(X)$, $\mu(X) = 0 \implies I(\mu) \leq 0$.
3. For all $\mu, \nu \in \mathcal{M}(X)$, $\mu(X) = \nu(X) \implies I(\mu) + I(\nu) \leq 2I(\mu, \nu)$.
4. For all $\mu, \nu \in \mathcal{M}^+(X)$, $\mu(X) = \nu(X) \implies I(\mu) + I(\nu) \leq 2I(\mu, \nu)$.
5. For all $\mu, \nu \in \mathcal{M}^1(X)$, $I(\mu) + I(\nu) \leq 2I(\mu, \nu)$.
6. I is concave on $\{\mu \in \mathcal{M}(X) : \mu(X) = 1\}$.

Proof. To show that (1) \iff (2), suppose that X is quasihypermetric. Let d denote the metric of X and let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$. If μ is atomic then we can write $\mu = \sum_{i=1}^n w_i \delta_{x_i}$ for some $w_1, \dots, w_n \in \mathbb{R}$ and some $x_1, \dots, x_n \in X$. By Theorem 5.2.6 it follows that

$$\mu(X) = \sum_{i=1}^n w_i = 0 \implies I(\mu) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0.$$

Otherwise, let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of μ , and by Theorem 2.3.24 there exist sequences of atomic measures $(\mu_n^+), (\mu_n^-) \in \mathcal{M}^+(X)$ such that $\mu_n^+ \rightarrow^* \mu^+$ and $\mu_n^- \rightarrow^* \mu^-$ and for each n , $\mu_n^+(X) = \mu^+(X)$ and $\mu_n^-(X) = \mu^-(X)$. Let $(\mu_n) \in \mathcal{M}(X)$ be the sequence defined for $n \in \mathbb{N}$ by $\mu_n = \mu_n^+ - \mu_n^-$. Then for each n ,

$$\|\mu_n\| \leq \|\mu_n^+\| + \|\mu_n^-\| = \|\mu^+\| + \|\mu^-\|,$$

and by Theorem 3.2.1 it follows that $I(\mu_n) \rightarrow I(\mu)$. Since each μ_n is atomic, we have that $I(\mu_n) \leq 0$, giving $I(\mu) \leq 0$.

Conversely, suppose that for all $\mu \in \mathcal{M}(X)$, $\mu(X) = 0 \implies I(\mu) \leq 0$. Let $n \in \mathbb{N}$, let $w_1, \dots, w_n \in \mathbb{R}$, let $x_1, \dots, x_n \in X$ and let $\mu = \sum_{i=1}^n w_i \delta_{x_i}$. Then

$$\sum_{i=1}^n w_i = \mu(X) = 0 \implies \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) = I(\mu) \leq 0,$$

and by Theorem 5.2.6 it follows that X is quasihypermetric.

To show that (2) \implies (3), suppose that for all $\mu \in \mathcal{M}(X)$, $\mu(X) = 0$ implies that $I(\mu) \leq 0$. Let $\mu, \nu \in \mathcal{M}(X)$ such that $\mu(X) = \nu(X)$. Then $(\mu - \nu)(X) = 0$ and $I(\mu - \nu) = I(\mu) + I(\nu) - 2I(\mu, \nu) \leq 0$, giving that $I(\mu) + I(\nu) \leq 2I(\mu, \nu)$.

It is clear that (3) \implies (4) \implies (5). To show that (5) \implies (2), suppose that for all $\mu \in \mathcal{M}^1(X)$, $I(\mu) + I(\nu) \leq 2I(\mu, \nu)$. Let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$. If $\mu = 0$ then the result is obvious. Otherwise, let $\mu = \mu^+ - \mu^-$. Then $\|\mu^+\| = \|\mu^-\| \neq 0$, and given that $\mu^+/\|\mu^+\|, \mu^-/\|\mu^-\| \in \mathcal{M}^1(X)$, we have that

$$\begin{aligned} I(\mu) &= I\left(\|\mu^+\| \cdot \frac{\mu^+}{\|\mu^+\|} - \|\mu^-\| \cdot \frac{\mu^-}{\|\mu^-\|}\right) \\ &= \|\mu^+\|^2 \cdot \left(I\left(\frac{\mu^+}{\|\mu^+\|}\right) + I\left(\frac{\mu^-}{\|\mu^-\|}\right) - 2I\left(\frac{\mu^+}{\|\mu^+\|}, \frac{\mu^-}{\|\mu^-\|}\right)\right) \\ &\leq \|\mu^+\|^2 \cdot 0 \\ &= 0. \end{aligned}$$

Finally, we will show that (3) \implies (6) and (6) \implies (2). Suppose that for all $\mu, \nu \in \mathcal{M}(X)$, $\mu(X) = \nu(X) \implies I(\mu) + I(\nu) \leq 2I(\mu, \nu)$. Let $\mu, \nu \in \mathcal{M}(X)$ such that $\mu(X) = \nu(X) = 1$ and let $\alpha \in [0, 1]$. We then have

$$\begin{aligned} &\alpha I(\mu) + (1 - \alpha)I(\nu) \\ &= \alpha(\alpha + 1 - \alpha)I(\mu) + (1 - \alpha)(\alpha + 1 - \alpha)I(\nu) \\ &= \alpha^2 I(\mu) + (1 - \alpha)^2 I(\nu) + \alpha(1 - \alpha)(I(\mu) + I(\nu)) \\ &\leq \alpha^2 I(\mu) + (1 - \alpha)^2 I(\nu) + 2\alpha(1 - \alpha)I(\mu, \nu) \\ &= I(\alpha\mu) + I((1 - \alpha)\nu) + 2I(\alpha\mu, (1 - \alpha)\nu) \\ &= I(\alpha\mu + (1 - \alpha)\nu); \end{aligned}$$

that is, I is concave on $\{\mu \in \mathcal{M}(X) : \mu(X) = 1\}$.

Suppose that I is concave $\{\mu \in \mathcal{M}(X) : \mu(X) = 1\}$. Let $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$ and let $x \in X$. Then $\mu + \delta_x, \delta_x \in \mathcal{M}(X)$ and $(\mu + \delta_x)(X) = \delta_x(X) = 1$.

Noting that $I(\delta_x) = 0$, we then have

$$\begin{aligned}
 \frac{1}{2}I(\mu) &= \frac{1}{2}I(\mu + \delta_x) + \frac{1}{2}I(\delta_x) - I(\mu + \delta_x, \delta_x) \\
 &\leq I\left(\frac{1}{2}\mu + \delta_x\right) - I(\mu, \delta_x) - I(\delta_x, \delta_x) \\
 &= \frac{1}{4}I(\mu) + I(\delta_x) + I(\mu, \delta_x) - I(\mu, \delta_x) \\
 &= \frac{1}{4}I(\mu).
 \end{aligned}$$

It follows that $I(\mu) \leq 0$. □

It is easily seen that characterisation 3 of hypermetric spaces from Theorem 5.1.2 and characterisation 2 of quasihypermetric spaces from Theorem 5.3.1 give rise to an alternative proof that the class of quasihypermetric spaces contains the class of hypermetric spaces. It simply needs to be noted that we can arbitrarily weak-* approximate a finite signed Borel measure μ on a quasihypermetric space X by a sequence of atomic measures (μ_n) such that each μ_n can be written using rational weights and $(\|\mu_n\|)$ is bounded, and that the functional $I: \mathcal{M}(X) \rightarrow \mathbb{R}$ is continuous on sets of bounded measure norm.

5.3.2 Metric embedding characterisations

We now develop characterisations of quasihypermetric spaces via embeddings in ℓ_2 .

Let (X, d) be a metric space, and consider the function $d^{1/2}: X \times X$ defined for $(x, y) \in X \times X$ by $d^{1/2}(x, y) = \sqrt{d(x, y)}$. Given that the triangle inequality holds for d , and that for all x, y, z ,

$$\begin{aligned}
 d(x, y) &\leq d(x, z) + d(z, y) \\
 \implies d(x, y) &\leq d(x, z) + 2d^{1/2}(x, z)d^{1/2}(z, y) + d(z, y) \\
 \implies (d^{1/2}(x, y))^2 &\leq (d^{1/2}(x, z) + d^{1/2}(z, y))^2 \\
 \implies d^{1/2}(x, y) &\leq d^{1/2}(x, z) + d^{1/2}(z, y),
 \end{aligned}$$

the triangle inequality must also hold for $d^{1/2}$. It is a routine exercise to show that $d^{1/2}$ satisfies the other properties of a metric, and hence $(X, d^{1/2})$ is a metric space.

Definition 5.3.2. Let (X, d) and Y be metric spaces. If the metric space $(X, d^{1/2})$ is isometrically embeddable in Y then we say that (X, d) is *1/2-embeddable* in Y .

We find it convenient to now introduce the strict quasihypermetric property for finite metric spaces. Recall that for $n \in \mathbb{N}$, if $(\{x_1, \dots, x_n\}, d)$ is a finite quasihypermetric metric space then for all $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 0$, Theorem 5.2.6 gives that

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) \leq 0.$$

We will consider quasihypermetric spaces where equality can only occur in the above when $w_i = 0$ for each i .

Definition 5.3.3. Let $n \in \mathbb{N}$, let $X = \{x_1, \dots, x_n\}$ be a finite set and let d be a metric on X . If for all $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 0$, we have that

$$\sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j) = 0 \implies w_1 = \dots = w_n = 0,$$

then (X, d) is a strictly quasihypermetric space.

A more general discussion of this property can be found in Section 5.4. For now, note that Theorem 5.4.3 states that finite subsets of Euclidean spaces are strictly quasihypermetric.

The remainder of this section builds up to our main metric embedding characterisation of quasihypermetric spaces, which is given in Theorem 5.3.17. This theorem states that a metric space is quasihypermetric if and only if it is 1/2-embeddable in the hypermetric space ℓ_2 , and that when each finite subset of the space is embeddable on a sphere such that each circumradius of a finite 1/2-embedding is bounded above by some constant then the space is embeddable on a sphere in ℓ_2 .

Definition 5.3.4. Let A be a symmetric $n \times n$ matrix with real entries. If $x^T A x \leq 0$ for all $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ such that $\sum_{i=1}^n x_i = 0$ then A is *almost-negative definite*. Further, if equality holds only when $x = 0$ then A is *strictly almost-negative definite*.

Let (X, d) be a metric space, let $n \in \mathbb{N}$, let $x_1, \dots, x_n \in X$ and let A be the $n \times n$ distance matrix given by $A_{ij} = d(x_i, x_j)$. Then A is a symmetric $n \times n$ matrix with real entries and for all $w = (w_1, \dots, w_n) \in \mathbb{R}^n$,

$$w^T A w = \sum_{i=1}^n w_i \left(\sum_{j=1}^n d(x_i, x_j) w_j \right) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j d(x_i, x_j).$$

It follows that:

Theorem 5.3.5. *Let X be a metric space. Then X is quasihypermetric if and only if each distance matrix obtained from a finite tuple of points in X is almost-negative definite. Further, when X is finite, X is strictly quasihypermetric if and only if each distance matrix obtained from the list of distinct points in X is strictly almost-negative definite.*

The proofs of the following Lemmas and Theorem are due to Baxter [7, pg.4-5], whose derivation follows that of Schoenberg [46, pg.724-726].

Lemma 5.3.6. *Let $n \in \mathbb{N}$ and let A be a real non-zero symmetric positive semi-definite $n \times n$ matrix. Then there exists a real $n \times n$ matrix X such that $A = X X^T$, and by letting $x_1, \dots, x_n \in \mathbb{R}^n$ such that $\sqrt{2}x_1, \dots, \sqrt{2}x_n$ are the rows of X , for all i and j ,*

$$A_{ij} = \|x_i\|^2 + \|x_j\|^2 - \|x_i - x_j\|^2.$$

Further, if A is positive definite then X is non-singular.

Proof. As A is non-zero and symmetric, by the Spectral Theorem [49, pg. 222] it has n eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ and n associated linearly independent eigenvectors $p_1, \dots, p_n \in \mathbb{R}^n$, which may be chosen to be orthonormal. Given that A is positive semi-definite, then for each i , $\langle p_i, A p_i \rangle = \langle p_i, \lambda_i p_i \rangle = \lambda_i \|p_i\|^2 \geq 0$, and as $\|p_i\| \neq 0$, we have that $\lambda_i \geq 0$. Let D be the $n \times n$ diagonal matrix with diagonal entries $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$, let P be the $n \times n$ matrix with columns p_1, \dots, p_n and let $X = P D$. Then

$$A = P D^2 P^T = P D D^T P^T = (P D) (P D)^T = X X^T.$$

Now, let $x_1, \dots, x_n \in \mathbb{R}^n$ be such that $\sqrt{2}x_1, \dots, \sqrt{2}x_n$ are the rows of X . Then for all i and j ,

$$A_{ij} = \left\langle \sqrt{2}x_i, \sqrt{2}x_j \right\rangle = 2 \langle x_i, x_j \rangle = \|x_i\|^2 + \|x_j\|^2 - \|x_i - x_j\|^2.$$

Finally, if A is positive definite then $\lambda_i > 0$ for each i , which gives that the $\sqrt{\lambda_i}p_i$ are linearly independent, and so X is non-singular. \square

Lemma 5.3.7. *Let $n \in \mathbb{N}$ and let A be a real symmetric $n \times n$ matrix with diagonal entries zero. Then A is almost-negative definite if and only if the $n \times n$ matrix $(A_{ij} - A_{in} - A_{nj})_{ij}$ is negative semi-definite. Further, A is strictly almost-negative definite if and only if $(A_{ij} - A_{in} - A_{nj})_{ij}$ is negative definite.*

Proof. Suppose that A is almost-negative definite. Let $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. For each i and j , $x_i x_n (A_{in} - A_{in} - A_{nn}) = 0$ and $x_n x_j (A_{nj} - A_{nn} - A_{nj}) = 0$. Hence the value of

$$\sum_{i=1}^n \sum_{j=1}^n x_i x_j (A_{ij} - A_{in} - A_{nj})$$

is independent of x_n , and we may assume without loss of generality that

$$x_n = \sum_{i=1}^{n-1} -x_i.$$

Then

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n x_i x_j (A_{ij} - A_{in} - A_{nj}) \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} x_i x_j (A_{ij} - A_{in} - A_{nj}) \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} x_i x_j A_{ij} + \sum_{j=1}^{n-1} \left(\sum_{i=1}^{n-1} -x_i \right) x_j A_{in} + \sum_{i=1}^{n-1} x_i \left(\sum_{j=1}^{n-1} -x_j \right) A_{nj} \\ &= \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} x_i x_j A_{ij} + \sum_{j=1}^{n-1} x_n x_j A_{in} + \sum_{i=1}^{n-1} x_i x_n A_{nj} + x_n x_n A_{nn} \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i x_j A_{ij} \\ &\leq 0. \end{aligned}$$

Therefore $(A_{ij} - A_{in} - A_{nj})_{ij}$ is negative semi-definite. Further, it is easily seen that if A is strictly almost-negative definite then $(A_{ij} - A_{in} - A_{nj})_{ij}$ is negative definite.

Conversely, suppose that $(A_{ij} - A_{in} - A_{nj})_{ij}$ is negative semi-definite (resp. negative definite). The above calculation then shows that A is almost-negative definite (resp. strictly almost-negative definite). \square

Theorem 5.3.8. *Let $n \in \mathbb{N}$ and let A be a real non-zero $n \times n$ matrix with diagonal entries zero. Then A is (strictly) almost-negative definite if and only if there exist (affinely independent) $x_1, \dots, x_n \in \mathbb{R}^{n-1}$ such that for each i and j , $A_{ij} = \|x_i - x_j\|^2$.*

Proof. Suppose that A is almost-negative definite. Let e_1, \dots, e_n be the standard basis for \mathbb{R}^n , let $b_1, \dots, b_n \in \mathbb{R}^n$ be defined for each i by $b_i = e_i - e_n$, let B be the $n \times n$ matrix with column vectors b_1, \dots, b_n and let $C = -B^T A B$. It is easily seen that $C_{ij} = -\langle b_i, A b_j \rangle$. Let $a_1, \dots, a_n \in \mathbb{R}^n$ be the columns of A . Then for each i and j , $A e_i = a_i$ and $\langle e_i, a_j \rangle = A_{ij}$ and

$$\begin{aligned} C_{ij} &= -\langle e_i - e_n, A(e_j - e_n) \rangle \\ &= -\langle e_i - e_n, a_j - a_n \rangle \\ &= -\langle e_i, a_j \rangle + \langle e_i, a_n \rangle + \langle e_n, a_j \rangle - \langle e_n, a_n \rangle \\ &= -A_{ij} + A_{in} + A_{nj}. \end{aligned}$$

Using Lemma 5.3.7, C is positive semi-definite, and hence by Lemma 5.3.6 there exist $y_1, \dots, y_n \in \mathbb{R}^n$ such that $C_{ij} = \|y_i\|^2 + \|y_j\|^2 - \|y_i - y_j\|^2$ for each i and j . Noting that $j = i$ implies $A_{in} = \|y_i\|^2$, it follows that $A_{ij} = \|y_i - y_j\|^2$ for each i and j . Let $x_1, \dots, x_n \in \mathbb{R}^n$ be defined for each i by $x_i = y_i - y_n$. Then for each i and j , $A_{ij} = \|x_i - x_j\|^2$. Using a suitable rotation, we may consider that $x_i \in \mathbb{R}^{n-1}$. Further, if A is strictly almost-negative definite then by Lemma 5.3.6 the y_i are linearly independent and hence the x_i are affinely independent.

Conversely, suppose that there exist $x_1, \dots, x_n \in \mathbb{R}^{n-1}$ such that for each i and j , $A_{ij} = \|x_i - x_j\|^2$. Then A is symmetric and for all $w = (w_1, \dots, w_n) \in \mathbb{R}^n$

such that $\sum_{i=1}^n w_i = 0$,

$$\begin{aligned}
w^T A w &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j \|x_i - x_j\|^2 \\
&= \sum_{i=1}^n \sum_{j=1}^n w_i w_j (\|x_i\|^2 + \|x_j\|^2 - 2 \langle x_i, x_j \rangle) \\
&= 2 \sum_{i=1}^n \left(w_i \|x_i\|^2 \sum_{j=1}^n w_j \right) - 2 \sum_{i=1}^n \sum_{j=1}^n w_i w_j \langle x_i, x_j \rangle \\
&= -2 \left\langle \sum_{i=1}^n w_i x_i, \sum_{j=1}^n w_j x_j \right\rangle \\
&= -2 \left\| \sum_{i=1}^n w_i x_i \right\|^2 \\
&\leq 0.
\end{aligned}$$

Therefore, A is almost-negative definite. Further, suppose that the x_i are affinely independent, and let $w_1, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=1}^n w_i = 0$ and $w^T A w = 0$. By considering the above, we must then have that $\sum_{i=1}^n w_i x_i = 0$, and by the affine independence of the x_i it follows that $w_i = 0$ for each i . Therefore A is strictly almost-negative definite. \square

It is then immediate that:

Theorem 5.3.9. *Let X be a metric space. Then X is quasihypermetric if and only if for each $n \in \mathbb{N}_0$, each finite subset of X with $n + 1$ points can be $1/2$ -embedded in \mathbb{R}^n . Further, when X is a finite space containing $n + 1$ points for some $n \in \mathbb{N}_0$, X is strictly quasihypermetric if and only if X can be affinely $1/2$ -embedded in \mathbb{R}^n .*

We now see that for finite strictly quasihypermetric spaces, the $1/2$ -embedding given by the previous theorem is spherical. The following result is well-known.

Lemma 5.3.10. *Let $n \in \mathbb{N}$ and let $x_0, \dots, x_n \in \mathbb{R}^n$ be affinely independent. Then there exists a unique sphere in \mathbb{R}^n containing the x_i .*

Proof. Suppose that $x \in \mathbb{R}$ is the centre of a sphere containing the x_i . Then for each i and j , x is equidistant from x_i and x_j , and must lie on their perpendicular bisector, giving

$$\left\langle x - \frac{x_i + x_j}{2}, x_i - x_j \right\rangle = 0,$$

and hence $\langle x_i - x_j, x \rangle = (1/2)(\|x_i\|^2 - \|x_j\|^2)$. To show that there exists a unique sphere containing the x_i , we want to show that the system of equations

$$\langle x_i - x_j, x \rangle = (1/2)(\|x_i\|^2 - \|x_j\|^2) \quad \text{for } i, j = 0, \dots, n$$

has a unique solution for the unknown x . As this system is equivalent to

$$\langle x_i - x_{i+1}, x \rangle = (1/2)(\|x_i\|^2 - \|x_{i+1}\|^2) \quad \text{for } i = 0, \dots, n-1,$$

it will be sufficient to show that the vectors $x_i - x_{i+1}$ for $i = 0, \dots, n-1$ are linearly independent.

Let $c_0, \dots, c_{n-1} \in \mathbb{R}$ such that $\sum_{i=0}^{n-1} c_i(x_i - x_{i+1}) = 0$. Then

$$c_0 x_0 + \sum_{i=1}^{n-1} (c_i - c_{i-1}) x_i - c_{n-1} x_n = 0,$$

and by adding the coefficients of the x_i as a finite telescoping series, we obtain

$$c_0 + \sum_{i=1}^{n-1} (c_i - c_{i-1}) - c_{n-1} = 0.$$

As the x_i are affinely independent, it follows that $c_i = 0$ for each i , and hence the $x_i - x_{i+1}$ for $i = 0, \dots, n-1$ are linearly independent. \square

We then obtain:

Theorem 5.3.11. *Let X be a finite metric space with $n+1$ points for some $n \in \mathbb{N}_0$. If X is strictly quasihypermetric then X can be $1/2$ -embedded on a sphere in \mathbb{R}^n . Further, X cannot be $1/2$ -embedded on a sphere in \mathbb{R}^m when $m < n$, and the circumradius of the $1/2$ -embedding of X in \mathbb{R}^n is less than or equal to the circumradius of any $1/2$ -embedding of X in \mathbb{R}^m when $m \geq n$.*

Note that $1/2$ -embeddability on a sphere in a Euclidean space is not sufficient to ensure that a finite metric space is strictly quasihypermetric, as will be seen in comments given after Proposition 5.4.2.

Making a brief digression, the previous result allows us to give an example of a quasihypermetric space which is neither hypermetric nor strictly quasihypermetric. The example is due to Assouad [5, pg. 361-362]. (Assouad's paper does not give details of the weights which show that the space is not hypermetric, but refers the reader to his thesis written at the Université de Paris.)

Proposition 5.3.12. *There exists a five-point quasihypermetric space which is neither hypermetric nor strictly quasihypermetric.*

Proof. Let $X = \{1, 2, 3, 4, 5\}$, let A be the 5×5 matrix

$$A = \begin{pmatrix} 0 & 2 & 2 & 5 & 5 \\ 2 & 0 & 4 & 3 & 3 \\ 2 & 4 & 0 & 3 & 3 \\ 5 & 3 & 3 & 0 & 4 \\ 5 & 3 & 3 & 4 & 0 \end{pmatrix},$$

and let $d: X \times X \rightarrow \mathbb{R}$ be the function defined for $(i, j) \in X \times X$ by $d(i, j) = A_{ij}$. Given that A is a non-negative symmetric matrix with zeros on the main diagonal and nowhere else, to show that (X, d) is a metric space it will be sufficient to show that d satisfies the triangle inequality. Suppose that there exists distinct $i, j, k \in X$ such that $d(i, k) > d(i, j) + d(j, k)$. Given that $d(i, j) + d(j, k) \geq 4$, it follows that $d(i, k) = 5$, and we may assume without loss of generality that $i = 1$ and $k \in \{4, 5\}$. It can then be seen that $j \in \{2, 3\}$, giving that $d(i, j) + d(j, k) \leq 2 + 3 = 5 = d(i, k)$, which is a contradiction. Hence d satisfies the triangle inequality.

Now, since a distance matrix of the Euclidean space

$$\{(-1, 0, 0), (0, 1, 0), (0, -1, 0), (1, 0, 1), (1, 0, -1)\} \subseteq \mathbb{R}^3$$

obtained by using each point exactly once is

$$\begin{pmatrix} 0 & \sqrt{2} & \sqrt{2} & \sqrt{5} & \sqrt{5} \\ \sqrt{2} & 0 & 2 & \sqrt{3} & \sqrt{3} \\ \sqrt{2} & 2 & 0 & \sqrt{3} & \sqrt{3} \\ \sqrt{5} & \sqrt{3} & \sqrt{3} & 0 & 2 \\ \sqrt{5} & \sqrt{3} & \sqrt{3} & 2 & 0 \end{pmatrix},$$

it is easily seen that (X, d) can be $1/2$ -embedded in \mathbb{R}^3 , and by Theorem 5.3.9, (X, d) is quasihypermetric but not strictly quasihypermetric. To see that (X, d) is not hypermetric, let $w_1 = w_4 = w_5 = 1$ and $w_2 = w_3 = -1$. Then $w_i \in \mathbb{Z}$ for each i and $\sum_{i=1}^5 w_i = 1$ and

$$\sum_{i=1}^5 \sum_{j=1}^5 w_i w_j d(i, j) = 2(-2 - 2 + 5 + 5 + 4 - 3 - 3 - 3 - 3 + 4) = 4 > 0. \quad \square$$

We now extend our $1/2$ -embedding theorems from finite spaces to countable spaces and finally to separable spaces. Towards this aim, we require the result that an isometry between arbitrary subsets of \mathbb{R}^n can be extended to an isometry of the entire space. Blumenthal claims in [9, pg. 93] that this result is “either well known to the reader” or that there is a “proof [...] which [they] can readily supply”. The proof supplied here was not found in any reference, but is most probably not original.

Lemma 5.3.13. *Let $A, B \subseteq \mathbb{R}^n$ such that $0 \in A$ and $0 \in B$, and let $f: A \rightarrow B$ be an isometry such that $f(0) = 0$. Then f is inner product preserving.*

Proof. Recall that for all $x, y \in \mathbb{R}^n$, we have the identity

$$\langle x, y \rangle = \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x - y\|^2),$$

which then gives for all $x, y \in A$,

$$\begin{aligned} \langle f(x), f(y) \rangle &= \frac{1}{2} (\|f(x) - f(0)\|^2 + \|f(y) - f(0)\|^2 - \|f(x) - f(y)\|^2) \\ &= \frac{1}{2} (\|x - 0\|^2 + \|y - 0\|^2 - \|x - y\|^2) \\ &= \langle x, y \rangle. \end{aligned}$$

Hence f is inner product preserving. \square

Proposition 5.3.14. *Let $A, B \subseteq \mathbb{R}^n$, and let $f: A \rightarrow B$ be an isometry. Then there exists an affine isometry of \mathbb{R}^n which extends f .*

Proof. By considering suitable translations, we will assume that $0 \in A$ and $0 \in B$ and $f(0) = 0$ and show that there exists a linear isometry of \mathbb{R}^n which extends f .

Suppose first that A is a spanning set for \mathbb{R}^n . Let $a_1, \dots, a_n \in A$ be a basis for \mathbb{R}^n and let $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the linear transformation such that for each i , $T(a_i) = f(a_i)$. We will firstly show that $T|_A = f$. Now, it is known [36, pg. 155,423] that for all $m \in \mathbb{N}$ and for all $v_1, \dots, v_m \in \mathbb{R}^n$, the determinant of the Gram matrix $(\langle v_i, v_j \rangle)_{ij}$ is non-negative and is positive if and only if the v_i are independent. It then follows from Lemma 5.3.13 that the matrices $(\langle a_i, a_j \rangle)_{ij} = (\langle f(a_i), f(a_j) \rangle)_{ij}$ are non-singular, and hence that the $f(a_i)$ are independent. Subsequently, let $x \in A$ and write x and $f(x)$ as the unique linear combinations

$$x = \sum_{i=1}^n \alpha_i a_i \quad \text{and} \quad f(x) = \sum_{i=1}^n \beta_i f(a_i).$$

Since $\langle f(a_i), f(x) \rangle = \langle a_i, x \rangle$ for each i , we then have

$$\left\langle f(a_i), \sum_{j=1}^n \beta_j f(a_j) \right\rangle = \left\langle a_i, \sum_{j=1}^n \alpha_j a_j \right\rangle,$$

giving

$$\sum_{j=1}^n \beta_j \langle a_i, a_j \rangle = \sum_{j=1}^n \alpha_j \langle a_i, a_j \rangle \quad \text{for } i = 1, \dots, n.$$

Recalling that the matrix $(\langle a_i, a_j \rangle)_{ij}$ is non-singular, it follows that $\alpha_i = \beta_i$ for each i . Then since $T(x) = \sum_{i=1}^n \alpha_i f(a_i)$, we must have $T(x) = f(x)$ for all $x \in A$; that is,

T is a linear transformation of \mathbb{R}^n which extends f . The calculation

$$\begin{aligned}
 \langle T(x), T(y) \rangle &= \left\langle T \left(\sum_{i=1}^n \alpha_i a_i \right), T \left(\sum_{j=1}^n \beta_j a_j \right) \right\rangle \\
 &= \left\langle \sum_{i=1}^n \alpha_i f(a_i), \sum_{j=1}^n \beta_j f(a_j) \right\rangle \\
 &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \langle f(a_i), f(a_j) \rangle \\
 &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j \langle a_i, a_j \rangle \\
 &= \left\langle \sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \beta_j a_j \right\rangle \\
 &= \langle x, y \rangle.
 \end{aligned}$$

which holds for all $x = \sum_{i=1}^n \alpha_i a_i, y = \sum_{i=1}^n \beta_i a_i \in \mathbb{R}^n$ shows that T is an isometry.

Otherwise, suppose that A is not a spanning set for \mathbb{R}^n . Let $m < n$ be the dimension of the space spanned by A , let $v_1, \dots, v_m \in A$ be independent, let $v'_1, \dots, v'_m \in \mathbb{R}^n$ such that $v'_i = f(v_i)$ for each $i = 1, \dots, m$, and let $v_{m+1}, \dots, v_n \in \mathbb{R}^n$ be such that v_1, \dots, v_n are independent. Now, consider the system of equations with unknown $x \in \mathbb{R}^n$ given by

$$\langle v'_i, x \rangle = \langle v_i, v_{m+1} \rangle \quad \text{for } i = 1, \dots, m.$$

Given that the v'_i must be independent, there exist solutions for x , determined uniquely up to $n - m$ parameters. Let v'_{m+1} be such a solution of the above system, and we have that $\langle v'_i, v'_j \rangle = \langle v_i, v_j \rangle$ for all $i, j = 1, \dots, m + 1$ and that v'_1, \dots, v'_{m+1} are independent. Continuing this process for a finite number of times, there exist independent v'_1, \dots, v'_n such that $\langle v'_i, v'_j \rangle = \langle v_i, v_j \rangle$ for all $i, j = 1, \dots, n$. Then by the above argument, the linear transformation of \mathbb{R}^n such that $v_i \mapsto v'_i$ for each i is a linear isometry which extends f . □

The following result concerning embeddings in \mathbb{R}^n is due to Menger. Alexander and Stolarsky claim that Blumenthal [9, Chapter 4] “thoroughly [discusses] results of

this nature”, but we were unable to find a proof for the case of spherical embeddings in his book. We did not obtain Menger’s papers (written in German), so we do not know how Menger’s proof proceeded in the case of spherical embeddings. The proof of the following result in the case of finite embeddings is based on that found in Blumenthal [9, pg. 132-133], and the proof for the case of spherical embeddings is original.

Theorem 5.3.15. *Let X be a countable metric space. Then X is embeddable in ℓ_2 if and only if for each $n \in \mathbb{N}_0$, each finite subset of X with $n+1$ points is embeddable in \mathbb{R}^n . Further, suppose that for each $n \in \mathbb{N}_0$, each finite subset A of X with $n+1$ points is embeddable on a sphere of minimal radius at most $r(A)$ in \mathbb{R}^n , and that $r = \sup r(A) < \infty$. Then X is embeddable on a sphere in ℓ_2 with radius r .*

Proof. It follows from Lemma 5.1.8 that if X is embeddable in ℓ_2 then for each $n \in \mathbb{N}_0$, each finite subset of X with $n+1$ points is embeddable in \mathbb{R}^n .

Suppose that for each $n \in \mathbb{N}_0$, each finite subset of X with $n+1$ points is embeddable in \mathbb{R}^n . If X is a finite space then we can obviously embed X in some \mathbb{R}^n , and subsequently in ℓ_2 . Otherwise, X is countably infinite. Let x_0, x_1, \dots be an enumeration of X and for each $n \geq 0$, let $X_n = \{x_0, x_1, \dots, x_n\}$. Also, for each $n, m \in \mathbb{N}$ such that $n \leq m$, let $T_{n,m}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $T_{n,\infty}: \mathbb{R}^n \rightarrow \ell_2$ be the lifting isometries defined for $(x_1, \dots, x_n) \in \mathbb{R}^n$ by

$$\begin{aligned} T_{n,m}(x_1, \dots, x_n) &= (x_1, \dots, x_n, 0, \dots, 0), \\ T_{n,\infty}(x_1, \dots, x_n) &= (x_1, \dots, x_n, 0, 0, \dots). \end{aligned}$$

We now construct a sequence $(f_n)_{n \geq 0}$ of isometries such that each f_n maps $X_n \rightarrow \mathbb{R}^n$. These isometries will have the property that if $n < m$ then $(T_{n,m} \circ f_n)(x_i) = f_m(x_i)$ for each $0 \leq i \leq n$.

Let $f_0: X_0 \rightarrow \mathbb{R}^0$ be the function defined by $f_0(x_0) = 0$. Then f_0 is an isometry. Assume that for some $n \in \mathbb{N}_0$ there exists an isometry $f_n: X_n \rightarrow \mathbb{R}^n$, and note that the relation on the class of metric spaces which associates isometric spaces is an equivalence relation. We then have that X_n is isometric to $\{f_n(x_i): i = 0, \dots, n\}$,

which is again isometric to $\{(T_{n,n+1} \circ f_n)(x_i) : i = 0, \dots, n\}$. Now, let $Y \subseteq \mathbb{R}^{n+1}$ be such that X_{n+1} is isometric to Y . Writing $Y = \{y_0, \dots, y_{n+1}\}$, we must have that $\{y_i : i = 0, \dots, n\}$ is isometric to $\{(T_{n,n+1} \circ f_n)(x_i) : i = 0, \dots, n\}$ and by Proposition 5.3.14, the isometry identifying these spaces can be extended to an isometry of \mathbb{R}^{n+1} . Hence there exists $y'_{n+1} \in \mathbb{R}^{n+1}$ such that $\{y_i : i = 0, \dots, n+1\}$ is isometric to $\{(T_{n,n+1} \circ f_n)(x_i) : i = 0, \dots, n\} \cup \{y'_{n+1}\}$. Let $f_{n+1} : X_{n+1} \rightarrow \mathbb{R}^{n+1}$ be the function defined for x_i by

$$f_{n+1}(x_i) = \begin{cases} (T_{n,n+1} \circ f_n)(x_i) & \text{if } 0 \leq i \leq n, \\ y'_{n+1} & \text{if } i = n+1, \end{cases}$$

which is clearly an isometry. Hence by induction, there exists a sequence $(f_n)_{n \geq 0}$ of isometries such that each f_n maps $X_n \rightarrow \mathbb{R}^n$. Further, let $n < m$ and let $0 \leq i \leq n$. Then by the definition of f_{n+1} ,

$$\begin{aligned} (T_{n,m} \circ f_n)(x_i) &= ((T_{n+1,m} \circ T_{n,n+1}) \circ f_n)(x_i) \\ &= (T_{n+1,m} \circ (T_{n,n+1} \circ f_n))(x_i) \\ &= (T_{n+1,m} \circ f_{n+1})(x_i). \end{aligned}$$

By finite induction on n , we then obtain

$$(T_{n,m} \circ f_n)(x_i) = (T_{m,m} \circ f_m)(x_i) = f_m(x_i).$$

Finally, let $f : X \rightarrow \ell_2$ be the function defined for x_i by $f(x_i) = (T_{i,\infty} \circ f_i)(x_i)$, and let $i < j$. Then we must have that

$$\begin{aligned} \|f(x_i) - f(x_j)\| &= \|(T_{i,\infty} \circ f_i)(x_i) - (T_{j,\infty} \circ f_j)(x_j)\| \\ &= \|((T_{j,\infty} \circ T_{i,j}) \circ f_i)(x_i) - (T_{j,\infty} \circ f_j)(x_j)\| \\ &= \|(T_{j,\infty} \circ (T_{i,j} \circ f_i))(x_i) - (T_{j,\infty} \circ f_j)(x_j)\| \\ &= \|(T_{i,j} \circ f_i)(x_i) - f_j(x_j)\| \\ &= \|f_j(x_i) - f_j(x_j)\| \\ &= \|x_i - x_j\|. \end{aligned}$$

Hence f is an isometry, and therefore X is embeddable in ℓ_2 .

Further, suppose that the embedding $f_n(X_n)$ of each X_n in \mathbb{R}^n is contained in some sphere. We now show that there exists a sequence of spheres $(S_n)_{n \geq 0}$ such that for each n , $f_n(X_n) \subseteq S_n$ and S_n is of minimal radius and $T_{n,n+1}(S_n) \subseteq S_{n+1}$. Let $(k_n)_{n \geq 0}$ be a sequence such that for each n , k_n is the maximal number of affinely independent points in $f_n(X_n)$, and let $(X'_n)_{n \geq 0}$ be a sequence defined recursively by $X'_0 = f_0(X_0)$, and for each $n \geq 0$,

$$X'_{n+1} = \begin{cases} T_{n,n+1}(X'_n) & \text{if } k_{n+1} = k_n, \\ T_{n,n+1}(X'_n) \cup \{f_{n+1}(x_{n+1})\} & \text{if } k_{n+1} = k_n + 1. \end{cases}$$

Noting that for each n , $T_{n,n+1}(X'_n) \subseteq f_{n+1}(X_{n+1})$ and that isometries between linear spaces preserve affine independence, it can be shown that X'_n is an affinely independent subset of $f_n(X_n)$ with k_n points. Let $(S_n)_{n \geq 0}$ be a sequence of spheres defined for $n \geq 0$ by taking a sphere containing $f_n(X_n)$ and intersecting this with the affine hyperplane spanned by X'_n (this “slicing” operation results in a sphere). Now, the image of a set S under an isometry is a slice of a sphere if and only if S is slice of a sphere. Since each S_n can be rotated to be a sphere in some $(k_n - 1)$ -dimensional Euclidean space, it follows from Lemma 5.3.10 that these spheres have the minimal radius of any sphere containing X'_n . Proceeding by induction, we now show that $f_n(X_n) \subseteq S_n$ for each n .

Clearly as any set containing one or two points is affinely independent,

$$f_0(X_0) = X'_0 \subseteq S_0 \quad \text{and} \quad f_1(X_1) = X'_1 \subseteq S_1.$$

Suppose now that for some $n \geq 0$, $f_n(X_n) \subseteq S_n$. Then

$$(T_{n,n+1} \circ f_n)(X_n) = T_{n,n+1}(f_n(X_n)) \subseteq T_{n,n+1}(S_n).$$

Also, by intersecting S_{n+1} with the $(k_{n+1} - 1)$ -dimensional affine hyperplane spanned by X'_{n+1} , we obtain a sphere containing $T_{n,n+1}(X'_n)$, which by the rotation argument above and Lemma 5.3.10 must be $T_{n,n+1}(S_n)$. Hence $T_{n,n+1}(S_n) \subseteq S_{n+1}$. Writing

$$f_{n+1}(X_{n+1}) = (T_{n,n+1} \circ f_n)(X_n) \cup \{f_{n+1}(x_{n+1})\},$$

it will be sufficient to establish that $f_{n+1}(x_{n+1}) \subseteq S_{n+1}$. If $k_{n+1} = k_n + 1$, then $f_{n+1}(x_{n+1}) \in X'_{n+1} \subseteq S_{n+1}$. Otherwise, $k_{n+1} = k_n$. Then $f_{n+1}(x_{n+1})$ is contained in the $(k_n - 1)$ -dimensional affine hyperplane spanned by $X'_{n+1} = T_{n,n+1}(X'_n)$, and so as $X'_n \subseteq S_n$, $f_{n+1}(x_{n+1}) \subseteq T_{n,n+1}(X'_n) \subseteq T_{n,n+1}(S_n) \subseteq S_{n+1}$. Hence by induction, $f_n(X_n) \subseteq S_n$ for each n . Further the above induction proof showed that for $k \geq 0$, $f_k(X_k) \subseteq S_k$ implies $T_{k,k+1}(S_k) \subseteq S_{k+1}$. Hence $T_{n,n+1}(S_n) \subseteq S_{n+1}$ for each n , establishing all of our required properties of the sequence (S_n) .

Let $(r_n)_{n \geq 0}$ be a sequence such that each r_n is the radius of S_n . Then (r_n) is non-decreasing and bounded above by r . Further, the Monotone Convergence Theorem gives that r is the limit of (r_n) . Let $(c_n)_{n \geq 0}$ be the sequence such that for each n , c_n is the centre of S_n . We now consider the sequence of sphere slices $(T_{n,\infty}(S_n))_{n \geq 0}$ and the corresponding sequence of centres $(T_{n,\infty}(c_n))_{n \geq 0}$. For each $m, n \in \mathbb{N}$, if $T_{m,\infty}(c_m) = T_{n,\infty}(c_n)$ then $\|T_{m,\infty}(c_m) - T_{n,\infty}(c_n)\| = 0$. Otherwise, assume without loss of generality that $m < n$ and consider the triangle with vertices $T_{0,\infty}(x_0) = 0$, $T_{m,\infty}(c_m)$ and $T_{n,\infty}(c_n)$. The line segment joining $T_{m,\infty}(c_m)$ and $T_{n,\infty}(c_n)$ must be perpendicular to the m -dimensional plane containing $T_{m,\infty}(S_m)$, and thus the triangle has a right-angle at the vertex given by $T_{m,\infty}(c_m)$. By an application of Pythagoras' Theorem, it follows that $\|T_{m,\infty}(c_m) - T_{n,\infty}(c_n)\| = \sqrt{r_n^2 - r_m^2} \rightarrow 0$ as $m, n \rightarrow \infty$. Therefore $(T_{n,\infty}(c_n))$ is a Cauchy sequence, and given that ℓ_2 is complete, it must also be convergent. Let c be the limit of $(T_{n,\infty}(c_n))$. Then using continuity properties of the norm, we have that for each n , and for k no less than n ,

$$\begin{aligned}
\|f(x_n) - c\| &= \left\| (T_{n,\infty} \circ f_n)(x_n) - \lim_{k \rightarrow \infty} T_{k,\infty}(c_k) \right\| \\
&= \lim_{k \rightarrow \infty} \|((T_{k,\infty} \circ T_{n,k}) \circ f_n)(x_n) - T_{k,\infty}(c_k)\| \\
&= \lim_{k \rightarrow \infty} \|(T_{k,\infty} \circ (T_{n,k} \circ f_n))(x_n) - T_{k,\infty}(c_k)\| \\
&= \lim_{k \rightarrow \infty} \|(T_{n,k} \circ f_n)(x_n) - c_k\| \\
&= \lim_{k \rightarrow \infty} \|f_k(x_n) - c_k\| \\
&= \lim_{k \rightarrow \infty} r_k \\
&= r.
\end{aligned}$$

Therefore, c is the centre of a sphere of radius r containing $f(X)$. \square

Using a standard argument, we can now extend the result to separable spaces.

Theorem 5.3.16. *Let X be a separable metric space. Then X is embeddable in ℓ_2 if and only if for each $n \in \mathbb{N}_0$, each finite subset of X with $n + 1$ points is embeddable in \mathbb{R}^n . Further, suppose that for each $n \in \mathbb{N}_0$, each finite subset A of X with $n + 1$ points is embeddable on a sphere of radius at most $r(A)$ in \mathbb{R}^n , and that $r = \sup r(A) < \infty$. Then X is embeddable on a sphere in ℓ_2 with radius r .*

Proof. As Theorem 5.3.15 considered the case of X being countable, we will assume that X is uncountable. It follows from Lemma 5.1.8 that if X is embeddable in ℓ_2 then for each $n \in \mathbb{N}_0$, each finite subset of X with $n + 1$ points is embeddable in \mathbb{R}^n . To show the converse result, suppose that each finite subset of X with $n + 1$ points can be embedded in \mathbb{R}^n . Let $A = \{a_1, \dots\}$ be a countable dense subset of X . Then by Theorem 5.3.15, there exists an isometry $f: A \rightarrow \ell_2$. We now want to extend this isometry to X . Let $x \in X$, let $(x_n) \in A$ be a sequence such that $x_n \rightarrow x$ and let $g: X \rightarrow \ell_2$ be the function defined by

$$g(x) = \lim_{n \rightarrow \infty} f(x_n).$$

We need to show that the definition of g is independent of the particular sequence chosen to approximate points from $X \setminus A$. Let $(y_n) \in A$ be a convergent sequence with limit x . Then (x_n) and (y_n) are Cauchy, and as f is an isometry, $(f(x_n))$ and $(f(y_n))$ are also Cauchy. By the completeness of ℓ_2 , there exist $u, v \in \ell_2$ such that $\lim f(x_n) = u$ and $\lim f(y_n) = v$. Given that the sequence $x_1, y_1, x_2, y_2, \dots$ is Cauchy, there exists $w \in \ell_2$ such that $f(x_1), f(y_1), f(x_2), f(y_2), \dots$ has limit w . Then $w = u = v$, as u and v are the limits of subsequences of a sequence with limit w . Therefore

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} f(y_n).$$

Hence the expression used to define $g(x)$ is independent of the sequence chosen to approximate x . It is clear that $g|_A = f$.

Let d be the metric of X and let $x, y \in X$. Let $(x_n), (y_n) \in A$ be sequences such that $x_n \rightarrow x$ and $y_n \rightarrow y$. Then as f is an isometry,

$$\begin{aligned}
 \|g(x) - g(y)\| &= \left\| \lim_{n \rightarrow \infty} f(x_n) - \lim_{n \rightarrow \infty} f(y_n) \right\| \\
 &= \left\| \lim_{n \rightarrow \infty} (f(x_n) - f(y_n)) \right\| \\
 &= \lim_{n \rightarrow \infty} \|f(x_n) - f(y_n)\| \\
 &= \lim_{n \rightarrow \infty} \|x_n - y_n\| \\
 &= \|x - y\|.
 \end{aligned}$$

Therefore g is an isometry, and so X can be embedded in ℓ_2 .

Further, suppose that the embedding of each finite subset of X with $n + 1$ points in \mathbb{R}^n can be chosen to be on a sphere of radius at most $r' > 0$ in ℓ_2 . By Theorem 5.3.15, there exists $c \in \ell_2$ and $r > 0$ such that c is the centre of a sphere of radius r containing $f(A)$ and $r \leq r'$. Let $x \in X$ and let $(x_n) \in A$ be a sequence such that $x_n \rightarrow x$. Then

$$\|g(x) - c\| = \left\| \lim_{n \rightarrow \infty} g(x_n) - c \right\| = \lim_{n \rightarrow \infty} \|g(x_n) - c\| = \lim_{n \rightarrow \infty} r = r.$$

Therefore, X can also be embedded on the sphere centred at c of radius r . \square

We now arrive at our main result for this section:

Theorem 5.3.17. *Let X be a compact metric space. Then X is quasihypermetric if and only if X is $1/2$ -embeddable in ℓ_2 . For each finite strictly quasihypermetric subset A of X , let $r(A)$ denote the circumradius of a $1/2$ -embedding of A in a Euclidean space of minimal dimension. If each finite subset A of X is strictly quasihypermetric and $r = \sup r(A) < \infty$ then X is embeddable on a sphere in ℓ_2 .*

5.4 Strictly quasihypermetric spaces

Definition 5.4.1. Let X be a quasihypermetric space. If for all compact subsets Y of X and for all $\mu \in \mathcal{M}(Y)$ such that $\mu(Y) = 0$, $I(\mu) = 0 \implies \mu = 0$ then X is

strictly quasihypermetric.

This definition of strictly quasihypermetric spaces is consistent with the one given for finite spaces in Definition 5.3.3. Note that a compact quasihypermetric space will be strictly quasihypermetric when $I(\mu) = 0 \implies \mu = 0$ for all $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$.

Proposition 5.4.2. *S^1 metrized by arc length is a compact hypermetric (and hence quasihypermetric) space which is not strictly quasihypermetric.*

Proof. For convenience, we will associate S^1 and its arc length metric with the space of complex numbers of modulus 1 metrized by arc length, denoted (\mathbb{T}, d) . Let $(x_n) \in \mathbb{T}$ be a sequence and let $\mathbb{T}_u = \{z \in \mathbb{T} : \text{Im}(z) \geq 0\}$. Then either \mathbb{T}_u or $\overline{\mathbb{T} \setminus \mathbb{T}_u}$ must contain an infinite number of distinct terms of (x_n) . Suppose without loss of generality that \mathbb{T}_u contains an infinite number of terms of (x_n) . As \mathbb{T}_u is isometric to the compact space $[0, \pi]$, \mathbb{T}_u must contain some convergent subsequence of (x_n) , which is of course convergent in \mathbb{T} . Hence (\mathbb{T}, d) is compact.

Let \mathcal{T} be the Borel σ -algebra on (\mathbb{T}, d) , let μ be the measure such that 2μ is the usual normalised measure on the measurable space $(\mathbb{T}, \mathcal{T})$, and let $f: \mathbb{T} \rightarrow \mathcal{T}$ be the function defined for $z \in \mathbb{T}$ by $f(z) = \{w \in \mathbb{T} : d(z, w) \leq \pi/2\}$. It can be seen that for all $z_1, z_2 \in \mathbb{T}$, $d(z_1, z_2) = \mu(f(z_1) \Delta f(z_2))$. Then by Proposition 5.1.10, (\mathbb{T}, d) is hypermetric, from which it follows by Theorem 5.2.2 that (\mathbb{T}, d) is quasihypermetric.

Let $z_1, z_2, z_3, z_4 \in \mathbb{T}$ be such that the pairs (z_1, z_3) and (z_2, z_4) are diametrically opposite points and let $\mu = \sum_{i=1}^4 (-1)^{i+1} \delta_{z_i} \in \mathcal{M}(\mathbb{T})$. It is clear that $\mu(X) = 0$ and $\mu \neq 0$. Note that for all $z \in \mathbb{T}$,

$$d(z_1, z) + d(z_3, z) = d(z_2, z) + d(z_4, z) = \pi.$$

It follows that

$$\begin{aligned} I(\mu) &= \int (d(z_1, z) - d(z_2, z) + d(z_3, z) - d(z_4, z)) d\mu(z) \\ &= \int 0 d\mu \\ &= 0. \end{aligned}$$

Hence (\mathbb{T}, d) is not strictly quasihypermetric. \square

The proof of the above result essentially constructs a four point quasihypermetric space which not strictly quasihypermetric. Using results from the previous section, it is easily seen that this four point space is quasihypermetric but not strictly quasihypermetric simply by $1/2$ -embedding the points on a square in \mathbb{R}^2 with side length $\sqrt{\pi/2}$. (This method however would not demonstrate that the space is hypermetric.)

We know that the class of quasihypermetric spaces contain the class of hypermetric spaces and the class of strictly quasihypermetric spaces. To see an example of a strictly quasihypermetric space which is not hypermetric, the reader is referred to [40]. This example, Proposition 5.4.2, and the following result of Schoenberg demonstrate that the class of hypermetric spaces and the class of strictly quasihypermetric spaces overlap, but that neither contains the other.

We state without proof the following result of Schoenberg [47, pg. 789-791].

Theorem 5.4.3. *A finite subset of a Euclidean space is strictly quasihypermetric.*

We also state without proof the following result of Björck [8, pg. 258-259], whose result is an extension of Schoenberg's.

Theorem 5.4.4. *For each $n \in \mathbb{N}$, \mathbb{R}^n is strictly quasihypermetric.*

Attempting to write a proof of this result to reproduce in this thesis has not met with any success. We note that Proposition 3.2.17 and Theorem 5.1.2 give the result for $[0, 1]$ up to atomic measures which may be written as a linear combination of point measures using rational weights. Björck's proof relies on an integral formula given by M. Riesz and extended by Frostman, which is apparently a deep result. Demonstrating that the integral formula holds is the hard part of the proof, and is done by Riesz and Frostman using Fourier transform techniques.

We conclude this section by noting that examining the proof of Theorem 5.3.1 gives the following characterisations of strictly quasihypermetric spaces:

Theorem 5.4.5. *Let X be a compact metric space. The following are equivalent:*

1. X is strictly quasihypermetric.
2. For all $\mu \in \mathcal{M}(X)$, $\mu(X) = 0 \implies I(\mu) \leq 0$, with equality occurring if and only if $\mu = 0$.
3. For all $\mu, \nu \in \mathcal{M}(X)$, $\mu(X) = \nu(X) \implies I(\mu) + I(\nu) \leq 2I(\mu, \nu)$, with equality occurring if and only if $\mu = \nu$.
4. For all $\mu, \nu \in \mathcal{M}^+(X)$, $\mu(X) = \nu(X) \implies I(\mu) + I(\nu) \leq 2I(\mu, \nu)$, with equality occurring if and only if $\mu = \nu$.
5. For all $\mu, \nu \in \mathcal{M}^1(X)$, $I(\mu) + I(\nu) \leq 2I(\mu, \nu)$, with equality occurring if and only if $\mu = \nu$.
6. I is strictly concave on $\{\mu \in \mathcal{M}(X) : \mu(X) = 1\}$.

5.5 Summary

This chapter introduced the hypermetric, quasihypermetric and strictly quasihypermetric inequalities. For each of these inequalities, we gave examples of spaces which satisfied it and examples of spaces which did not. We showed by elementary arguments that the spaces \mathbb{R}^n , L^1 , L^2 and ℓ_2 are hypermetric.

We demonstrated some relationships between these three classes of metric spaces, namely that all hypermetric spaces and all strictly quasihypermetric spaces are quasihypermetric, and that the classes of hypermetric and strictly quasihypermetric spaces overlap but that neither contains the other.

Considerable time was spent characterising hypermetric and quasihypermetric spaces using average distances, measures, and certain embeddings in ℓ_2 . These characterisations are important features of these spaces, and will prove useful in Chapter 6 for studying average distances in compact metric spaces.

Chapter 6

Results in quasihypermetric spaces

In Chapter 4, we characterised the average distances $m(X)$, $M(X)$ and $\overline{M}(X)$ using measures, gave properties of these constants and demonstrated various techniques to calculate their values for certain concrete spaces X . We now continue these discussions further in light of the quasihypermetric property of metric spaces.

6.1 Finiteness of $\overline{M}(X)$

The following result is due to Peter Nickolas [40], and justifies our previous claims that studying quasihypermetric spaces is essential to properly discuss $\overline{M}(X)$.

Theorem 6.1.1. *Let X be a compact metric space which is not quasihypermetric. Then $\overline{M}(X) = \infty$.*

Proof. By Theorem 5.3.1 there exists $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$ and $I(\mu) > 0$. Let $x \in X$, and consider the sequence of signed measures $(n\mu + \delta_x) \in \mathcal{M}(X)$. Then $(n\mu + \delta_x)(X) = 1$ for each $n \in \mathbb{N}$, and

$$I(n\mu + \delta_x) = n^2 I(\mu) + I(\delta_x) + 2nI(\mu, \delta_x) = n^2 I(\mu) + 2nd_\mu(x).$$

It follows that $I(n\mu + \delta_x) \rightarrow \infty$ as $n \rightarrow \infty$. Hence $\overline{M}(X) = \infty$. \square

In future, we will assume that X is a compact quasihypermetric space when

we discuss $\overline{M}(X)$. The next results due to Nickolas and Wolf [40] show that the converse of Theorem 6.1.1 does not hold.

Proposition 6.1.2. *Let (X, d) be a compact quasihypermetric space admitting a signed measure $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 0$ and $I(\mu) = 0$ and μ is d -invariant and $d_\mu \neq 0$. Then $\overline{M}(X, d) = \infty$.*

Proof. Let $c \in \mathbb{R}$ such that $d_\mu \equiv c$. Then $c \neq 0$, and as $-\mu$ is a signed Borel measure of total mass 0 on X and $I(-\mu) = I(\mu)$ and $d_{-\mu} = -d_\mu$, we may assume without loss of generality μ has been chosen such that $c > 0$. Let $x \in X$, and consider the sequence of signed measures $(n\mu + \delta_x) \in \mathcal{M}(X)$. Then $(n\mu + \delta_x)(X) = 1$ for each $n \in \mathbb{N}$, and

$$I(n\mu + \delta_x) = n^2 I(\mu) + I(\delta_x) + 2nI(\mu, \delta_x) = 2nc.$$

It follows that $I(n\mu + \delta_x) \rightarrow \infty$ as $n \rightarrow \infty$. Hence $\overline{M}(X) = \infty$. \square

Proposition 6.1.3. *There exists a five-point quasihypermetric space (X, d) such that $\overline{M}(X, d) = \infty$.*

Proof. Let $X = \{1, 2, 3, 4, 5\}$, let A be the 5×5 matrix

$$A = \begin{pmatrix} 0 & 2 & 2 & 5 & 5 \\ 2 & 0 & 4 & 3 & 3 \\ 2 & 4 & 0 & 3 & 3 \\ 5 & 3 & 3 & 0 & 4 \\ 5 & 3 & 3 & 4 & 0 \end{pmatrix},$$

and let $d: X \times X \rightarrow \mathbb{R}$ be the function defined for $(i, j) \in X \times X$ by $d(i, j) = A_{ij}$.

It is shown in the proof of Proposition 5.3.12 that d is a metric on X . Now, let $\mu = 2\delta_1 - 2\delta_2 - 2\delta_3 + \delta_4 + \delta_5 \in \mathcal{M}(X)$. Then $\mu(X) = 0$ and

$$\begin{aligned} \frac{I(\mu)}{2} &= 2 \cdot (-2) \cdot 2 + 2 \cdot (-2) \cdot 2 + 2 \cdot 1 \cdot 5 + 2 \cdot 1 \cdot 5 \\ &\quad + (-2) \cdot (-2) \cdot 4 + (-2) \cdot 1 \cdot 3 + (-2) \cdot 1 \cdot 3 \\ &\quad + (-2) \cdot 1 \cdot 3 + (-2) \cdot 1 \cdot 3 \\ &\quad + 1 \cdot 1 \cdot 4 \\ &= 0, \end{aligned}$$

which gives that $I(\mu) = 0$. Finally, it is a routine calculation to show that

$$\begin{pmatrix} 0 & 2 & 2 & 5 & 5 \\ 2 & 0 & 4 & 3 & 3 \\ 2 & 4 & 0 & 3 & 3 \\ 5 & 3 & 3 & 0 & 4 \\ 5 & 3 & 3 & 4 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ -2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix},$$

and so $d_\mu \equiv 2$. Hence by Proposition 6.1.2, $\overline{M}(X, d) = \infty$. \square

We now develop a sufficient condition to ensure that $M(X) < \infty$ for finite quasihypermetric spaces X , and show that $M(X) < \infty$ for compact Euclidean spaces X .

The following result is due to Alexander and Stolarsky [2, pg. 12].

Lemma 6.1.4. *Let $n \in \mathbb{N}$ and let $x_0, \dots, x_n \in \mathbb{R}^n$ be points on a sphere with radius $r > 0$. Then for all $w_0, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=0}^n w_i = 1$,*

$$\sum_{i=0}^n \sum_{j=0}^n w_i w_j \|x_i - x_j\|^2 \leq 2r^2.$$

Further, if the x_i are affinely independent then equality will occur for a unique choice of the w_i .

Proof. We may assume without loss of generality that x_i are chosen such that the origin is the centre of the sphere. Hence

$$\begin{aligned} \sum_{i=0}^n \sum_{j=0}^n w_i w_j \|x_i - x_j\|^2 &= \sum_{i=0}^n \sum_{j=0}^n w_i w_j (\|x_i\|^2 + \|x_j\|^2 - 2 \langle x_i, x_j \rangle) \\ &= \sum_{i=0}^n \sum_{j=0}^n w_i w_j (2r^2 - 2 \langle x_i, x_j \rangle) \\ &= 2r^2 \sum_{i=0}^n \left(w_i \sum_{j=0}^n w_j \right) - 2 \sum_{i=0}^n \sum_{j=0}^n w_i w_j \langle x_i, x_j \rangle \\ &= 2r^2 - 2 \left\langle \sum_{i=0}^n w_i x_i, \sum_{j=0}^n w_j x_j \right\rangle \end{aligned}$$

$$\begin{aligned}
&= 2r^2 - 2 \left\| \sum_{i=0}^n w_i x_i \right\|^2 \\
&\leq 2r^2,
\end{aligned}$$

with equality occurring if and only if $\sum_{i=0}^n w_i x_i = 0$. Now, suppose that the x_i are affinely independent. Then recalling that every point of \mathbb{R}^n can be written as a unique affine combination of $n+1$ affinely independent points, there exists a unique choice of the w_i such that

$$\sum_{i=0}^n w_i x_i = 0 \quad \text{and} \quad \sum_{i=0}^n w_i = 1.$$

Hence

$$\sum_{i=0}^n \sum_{j=0}^n w_i w_j \|x_i - x_j\|^2 = 2r^2$$

for a unique choice of the w_i . □

Recall that for $n \in \mathbb{N}_0$, by Theorem 5.3.9 every finite quasihypermetric space with $n+1$ points is $1/2$ -embeddable in \mathbb{R}^n . Further, when such a space is also strictly quasihypermetric, the embedding is affinely independent and hence spherical by Lemma 5.3.10. We then obtain:

Corollary 6.1.5. *Let X be a finite quasihypermetric space which is $1/2$ -embeddable on a sphere in a Euclidean space. Then $\overline{M}(X) < \infty$.*

Corollary 6.1.6. *Let X be a finite strictly quasihypermetric space and let r be the circumradius of the $1/2$ -embedding of X in a Euclidean space of minimal dimension. Then $\overline{M}(X) = 2r^2$ and there exists a unique \overline{M} -maximal measure on X .*

Proof. Since X is finite, each of its points are isolated. Hence all Borel measures on X are atomic, and the result follows. □

The following result is essentially due to Alexander and Stolarsky [2, pg. 14-15]. (Alexander and Stolarsky do not calculate an explicit bound, instead only showing that such a bound exists.)

Theorem 6.1.7. *Let $n \in \mathbb{N}$ such that $n \geq 2$ and let X be a compact subset of \mathbb{R}^n . Then*

$$\overline{M}(X) \leq \frac{\sqrt{\pi} D(X) \Gamma(n/2 + 1/2)}{2 \Gamma(n/2)} < \infty.$$

Proof. Let $\mu \in \mathcal{M}(X)$ be an atomic signed measure such that $\mu(X) = 1$, and write $\mu = \sum_{i=1}^m w_i \delta_{x_i}$ for some $m \in \mathbb{N}$, for some $x_1, \dots, x_m \in X$ and for some $w_1, \dots, w_m \in \mathbb{R}$ such that $\sum_{i=1}^m w_i = 1$. By Lemma 5.1.6 we then have that for each i and j ,

$$\|x_i - x_j\| = \frac{1}{2V_{n-1}} \int_{S^{n-1}} |\langle x, x_i - x_j \rangle| dx,$$

where dx refers to integration with respect to the usual surface measure on S^{n-1} and V_{n-1} denotes the volume of the unit ball in \mathbb{R}^{n-1} . For each $x \in S^{n-1}$, let $X'(x) = \{\langle x, x_i \rangle : i = 1, \dots, n\} \subseteq \mathbb{R}$. By the Cauchy-Schwarz inequality, it follows that for all $x \in S^{n-1}$ and for all i and j ,

$$|\langle x, x_i \rangle - \langle x, x_j \rangle| = |\langle x, x_i - x_j \rangle| \leq \|x\| \cdot \|x_i - x_j\| \leq 1 \cdot D(X) = D(X),$$

and so $D(X'(x)) \leq D(X)$. Using Proposition 1.2.10, we then have

$$\sum_{i=1}^m \sum_{j=1}^m w_i w_j |\langle x, x_i \rangle - \langle x, x_j \rangle| \leq I(X'(x)) = \frac{D(X'(x))}{2} \leq \frac{D(X)}{2}$$

for all $x \in S^{n-1}$, giving that

$$\begin{aligned} I(\mu) &= \sum_{i=1}^m \sum_{j=1}^m w_i w_j \|x_i - x_j\| \\ &= \sum_{i=1}^m \sum_{j=1}^m w_i w_j \left(\frac{1}{2V_{n-1}} \int_{S^{n-1}} |\langle x, x_i - x_j \rangle| dx \right) \\ &= \frac{1}{2V_{n-1}} \int_{S^{n-1}} \left(\sum_{i=1}^m \sum_{j=1}^m w_i w_j |\langle x, x_i \rangle - \langle x, x_j \rangle| \right) dx \\ &\leq \frac{1}{2V_{n-1}} \int_{S^{n-1}} \frac{D(X)}{2} dx \\ &= \frac{D(X)}{4V_{n-1}} \int_{S^{n-1}} 1 dx \\ &= \frac{D(X) S_{n-1}}{4V_{n-1}}, \end{aligned}$$

where S_{n-1} denotes the surface area of S^{n-1} . Given that $\overline{M}(X)$ may be considered as $\sup I(\mu)$, where μ ranges over the atomic measures of unit mass on X , it follows that

$$\overline{M}(X) \leq \frac{D(X) S_{n-1}}{4V_{n-1}} < \infty.$$

It remains to be shown that

$$\frac{S_{n-1}}{4V_{n-1}} = \frac{\sqrt{\pi} \Gamma(n/2 + 1/2)}{2 \Gamma(n/2)}.$$

It is known [18] that the volume and surface area of a sphere of unit radius in \mathbb{R}^n are respectively

$$\frac{\pi^{n/2}}{\Gamma(n/2 + 1)} \quad \text{and} \quad \frac{2\pi^{n/2}}{\Gamma(n/2)},$$

giving that

$$V_{n-1} = \frac{\pi^{(n-1)/2}}{\Gamma((n-1)/2 + 1)} = \frac{\pi^{(n-1)/2}}{\Gamma(n/2 + 1/2)} \quad \text{and} \quad S_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

Hence

$$\frac{S_{n-1}}{4V_{n-1}} = \frac{1}{4} \cdot \frac{2\pi^{n/2}}{\Gamma(n/2)} \cdot \frac{\Gamma(n/2 + 1/2)}{\pi^{(n-1)/2}} = \frac{\sqrt{\pi} \Gamma(n/2 + 1/2)}{2 \Gamma(n/2)}. \quad \square$$

For each $n \in \mathbb{N}$ such that $n \geq 2$, let M_n denote the bound given for $\overline{M}(X)$ by the previous result for compact subsets X of \mathbb{R}^n with unit diameter. We then calculate:

n	2	3	4	5	6
M_n	$\pi/4$	1	$3\pi/8$	$4/3$	$15\pi/32$

We note that when $n = 3$, it follows from Proposition 4.4.3 that this bound is sharp.

Together with Proposition 1.2.10, the previous result gives that:

Corollary 6.1.8. *Let X be a compact subset of some Euclidean space. Then $\overline{M}(X) < \infty$.*

6.2 Properties of $m(X)$, $M(X)$ and $\overline{M}(X)$

We now discuss some properties of $m(X)$, $M(X)$ and $\overline{M}(X)$, essentially building on the discussion in Sections 4.2-4.4.

6.2.1 M -maximal and \overline{M} -maximal measures

The following result extends a theorem of Björck [8, pg. 258], using an alternate proof.

Proposition 6.2.1. *Let X be a compact quasihypermetric space. Then the set of M -maximal measures on X is a non-empty convex subset of $\mathcal{M}^1(X)$, and hence must be a singleton or infinite. Further, if X is strictly quasihypermetric then this set is a singleton; that is, an M -maximal measure on X is unique.*

Proof. Recall that by Theorem 4.3.3 there exists an M -maximal measure on X .

Let $\mu_1, \mu_2 \in \mathcal{M}^1(X)$ be M -maximal measures on X . Then for all $\alpha \in [0, 1]$, Theorem 5.3.1 gives that $\alpha I(\mu_1) + (1 - \alpha)I(\mu_2) \leq I(\alpha\mu_1 + (1 - \alpha)\mu_2)$. Given that the left-hand side of the above inequality equals $M(X)$ and that $\mathcal{M}^1(X)$ is convex, then from the definition of $M(X)$ it follows that

$$M(X) = I(\alpha\mu_1 + (1 - \alpha)\mu_2)$$

and so $\alpha\mu_1 + (1 - \alpha)\mu_2 \in \mathcal{M}^1(X)$ is also an M -maximal measure on X .

Now, suppose that X is strictly quasihypermetric. Then by Theorem 5.4.5, we must have that $\mu_1 = \mu_2$, giving that an M -maximal measure on X is unique. \square

The original result of Björck is:

Corollary 6.2.2. *There exists a unique M -maximal measure on all compact subsets of a Euclidean space.*

A similar result holds concerning \overline{M} -maximal measures, but recall from Proposition 4.4.3 that there is not always an \overline{M} -maximal measure on a compact metric space.

Proposition 6.2.3. *Let X be a compact quasihypermetric space. Then the set of \overline{M} -maximal measures on X is a convex subset of $\mathcal{M}(X)$, and hence must be empty, a singleton or infinite. Further, if X is strictly quasihypermetric then this set is empty or a singleton; that is, if there exists an \overline{M} -maximal measure on X then it is unique.*

Proof. The proof is identical to that of Theorem 6.2.1, except we are not guaranteed that \overline{M} -maximal measures exist on X . \square

The following result is due to Nickolas [40].

Proposition 6.2.4. *Let (X, d) be a compact quasihypermetric space. If there exists a d -invariant measure $\mu \in \mathcal{M}^1(X)$ then μ is M -maximal. If there exists a d -invariant measure $\mu \in \mathcal{M}(X)$ such that $\mu(X) = 1$ then μ is \overline{M} -maximal.*

Proof. Suppose that there exists $\mu \in \mathcal{M}^1(X)$ such that $d_\mu \equiv c$ for some $c \in \mathbb{R}$. Then $I(\mu) = c$ and for all $\nu \in \mathcal{M}^1(X)$ we have that $I(\mu) + I(\nu) \leq 2I(\mu, \nu)$, giving $c + I(\nu) \leq 2c$ and hence $I(\nu) \leq c$. It follows that $I(\mu) = M(X, d)$; that is, μ is M -maximal. The corresponding statement for signed measures is proven by assuming μ and ν to be signed measures of total mass 1. \square

It follows immediately from the previous result that:

Corollary 6.2.5. *Let (X, d) be a compact quasihypermetric space. If there exists a d -invariant Borel probability measure on X then $m(X, d) = M(X, d) = \overline{M}(X, d)$.*

6.2.2 Relationships between $m(X)$, $M(X)$ and $\overline{M}(X)$

The following result is essentially due to Wolf [56, pg. 391, 396-397], but the proof is simpler.

Proposition 6.2.6. *Let (X, d) be a compact connected metric space. If $m(X, d)$ and $M(X, d)$ are equal then a max-min Borel probability measure on X is M -maximal and d -invariant.*

Proof. Let $\mu \in \mathcal{M}^1(X)$ be a max-min measure. Then for all $x \in X$,

$$M(X, d) = m(X, d) \leq d_\mu(x).$$

Now, by integrating the above with respect to μ , we obtain $M(X, d) \leq I(\mu)$ and hence $M(X, d) = I(\mu)$, which is of course the statement that μ is M -maximal.

To show that μ is d -invariant, it is sufficient to show that $d_\mu(x) \leq M(X, d)$ for all $x \in X$, which is immediate from Proposition 4.3.5. \square

The original result of Wolf is:

Corollary 6.2.7. *Let (X, d) be a compact connected metric space. If $m(X, d)$ and $M(X, d)$ are equal then there exists $\mu \in \mathcal{M}^1(X)$ such that μ is M -maximal and d -invariant.*

It also follows that:

Corollary 6.2.8. *Let X be a compact connected strictly quasihypermetric space. If $m(X, d) = M(X, d)$ then X admits a unique max-min Borel probability measure.*

We note that the converse of Proposition 6.2.6 is not true. The proof due to Wolf [56, pg. 391, 397] is omitted.

Proposition 6.2.9. *There exists a compact connected metric space (X, d) such that there exists a d -invariant Borel probability measure on X and $m(X, d) < M(X, d)$.*

The following result is due to Nickolas [40].

Theorem 6.2.10. *Let (X, d) be a compact connected quasihypermetric space. The following are equivalent:*

1. *Two of $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$ are equal.*
2. *All of $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$ are equal.*
3. *There exists a d -invariant Borel probability measure on X .*

Proof. To show that (1) \implies (2), suppose that two of $m(X, d)$, $M(X, d)$ and $\overline{M}(X, d)$ are equal. To show that $m(X, d) = M(X, d) = \overline{M}(X, d)$, we consider three cases.

Case 1: $m(X, d) = M(X, d)$. Let $\mu \in \mathcal{M}^1(X)$ be a max-min measure. Then by Proposition 6.2.6, μ is M -maximal and by Proposition 6.2.4 must also be \overline{M} -maximal. Hence $M(X, d) = \overline{M}(X, d) = I(\mu)$.

Case 2: $m(X, d) = \overline{M}(X, d)$. The required result follows simply by considering the inequality $m(X, d) \leq M(X, d) \leq \overline{M}(X, d)$.

Case 3: $M(X, d) = \overline{M}(X, d)$. Suppose that $m(X, d) < M(X, d)$, and let $\mu \in \mathcal{M}^1(X)$ be an M -maximal measure, which by Corollary 6.2.7 must not be d -invariant. Then d_μ is not constant. Recall that Proposition 4.3.5 gives that $d_\mu(x) \leq M(X, d)$ for all $x \in X$, with equality occurring when x belongs to the support of μ . Since X is connected and Proposition 3.1.1 gives that d_μ is continuous, using the Intermediate Value Theorem we have that $d_\mu(x) = M(X, d) - \varepsilon$ for some $x \in X$ and some $\varepsilon \in (0, M(X, d)/2)$. Let α be such that

$$0 < \alpha < \frac{2\varepsilon}{M(X, d) - 2\varepsilon}$$

and let $\nu = (1 + \alpha)\mu - \alpha\delta_x \in \mathcal{M}(X)$. Then $\nu(X) = 1$ and

$$\begin{aligned} I(\nu) &= (1 + \alpha)^2 I(\mu) + \alpha^2 I(\delta_x) - 2(1 + \alpha)\alpha I(\mu, \delta_x) \\ &= (1 + \alpha)^2 M(X, d) - 2(1 + \alpha)\alpha d_\mu(x) \\ &= (1 + \alpha)^2 M(X, d) - 2(1 + \alpha)\alpha (M(X, d) - \varepsilon) \\ &= (1 - \alpha^2)M(X, d) + 2(1 + \alpha)\alpha\varepsilon. \end{aligned}$$

Now, it can be seen by solving a quadratic inequality in α that

$$\begin{aligned} (1 - \alpha^2)M(X, d) + 2(1 + \alpha)\alpha\varepsilon > M(X, d) &\iff (-M(X, d) + 2\varepsilon)\alpha^2 + (2\varepsilon)\alpha > 0 \\ &\iff 0 < \alpha < \frac{2\varepsilon}{M(X, d) - 2\varepsilon}. \end{aligned}$$

Hence $\overline{M}(X, d) \geq I(\nu) > M(X, d)$, which is a contradiction. We must then have that $m(X, d) = M(X, d)$.

That (2) \implies (3) is given by Proposition 6.2.6 and that (3) \implies (1) is a consequence of Corollary 6.2.5. □

We note the following result of Yost [16, pg. 269].

Proposition 6.2.11. *Let (X, d) be a compact connected subset of \mathbb{R}^n which is not a line segment. If there exists a d -invariant Borel probability measure on X then no three points of X are colinear.*

We then obtain the following corollary:

Corollary 6.2.12. *Let X be a compact connected subset of \mathbb{R}^n which is not a line segment and contains three colinear points. Then $m(X) < M(X) < \overline{M}(X)$.*

6.2.3 Discontinuity of $\overline{M}: H(S) \rightarrow \mathbb{R}$

Recall from Section 4.2 that for any metric space S , $H(S)$ is the set of all non-empty compact subsets of S equipped with its Hausdorff metric. We have seen that the mappings $m: H(S) \rightarrow \mathbb{R}$ and $M: H(S) \rightarrow \mathbb{R}$ are continuous. Quoting an inequality of Alexander [1, pg. 318], we outline the proof that this is not necessarily the case for the mapping $\overline{M}: H(S) \rightarrow \mathbb{R}$. The Alexander inequality is derived using Archimedes' Theorem and integral geometry techniques which are not considered in this thesis.

Proposition 6.2.13. *The mapping $\overline{M}: H(\mathbb{R}^3) \rightarrow \mathbb{R}$ is discontinuous at S^2 .*

Proof. Let $X \in H(\mathbb{R}^3)$ be the sphere S^2 and let $(X_n)_{n \geq 2} \in H(\mathbb{R}^3)$ be the sequence such that for each $n \in \mathbb{N}$ where $n \geq 2$, X_n is a sphere of radius $1 - 1/n$ centred at the origin. Also, let $\mu \in \mathcal{M}(X)$ be the usual normalised surface measure on X and let $(\mu_n)_{n \geq 2}$ be the sequence such that for each $n \in \mathbb{N}$ where $n \geq 2$, μ_n is the usual normalised surface measure on X_n .

Now, clearly $X \cup X_n \rightarrow X$ in $H(\mathbb{R}^3)$, and we will see in Proposition 6.4.1 that $\overline{M}(X) = 4/3$. Consider the sequence of signed measures $(\nu_n)_{n \geq 2}$ of unit mass defined for each $n \in \mathbb{N}$ where $n \geq 2$ by $\nu_n = n\mu - (n-1)\mu_n \in \mathcal{M}(X \cup X_n)$. It is shown by Alexander [1, pg. 318] that $I(\nu_n) \geq 2(1 - 1/n)$. Noting that $X \cup X_n \subseteq B_3$, using Proposition 4.4.3 (which is also a result of Alexander from the same paper) we obtain

$$2 \left(1 - \frac{1}{n}\right) \leq I(\nu_n) \leq \overline{M}(X \cup X_n) \leq \overline{M}(B_3) = 2.$$

Hence $\overline{M}(X \cup X_n) \rightarrow 2$ and $\overline{M}: H(\mathbb{R}^3) \rightarrow \mathbb{R}$ is discontinuous at $X = S^2$. \square

6.3 Geometrical interpretations of $M(X)$ and $\overline{M}(X)$

Recall that by Theorem 5.3.11, we know that a finite strictly quasihypermetric space X is $1/2$ -embeddable on a sphere in some Euclidean space of minimal dimension.

We now show that $M(X)$ and $\overline{M}(X)$ have a geometrical interpretation concerning the circumradius of such a $1/2$ -embedding and the distance from the circumcentre of the $1/2$ -embedding to the convex hull of the $1/2$ -embedding.

The following result is due to Alexander and Stolarsky [2, pg. 14-15].

Theorem 6.3.1. *Let (X, d) be a finite strictly quasihypermetric space, let r be the circumradius of a $1/2$ -embedding of X in some Euclidean space of minimal dimension and let s be the distance from the circumcentre of the $1/2$ -embedding to the convex hull of the $1/2$ -embedding. Then*

$$M(X, d) = 2(r^2 - s^2) \quad \text{and} \quad \overline{M}(X, d) = 2r^2,$$

and there exist unique M -maximal and \overline{M} -maximal measures.

Proof. The result clearly holds when X is a singleton. Otherwise, let $n \in \mathbb{N}_0$ such that X has $n + 1$ points and write $X = \{x_0, \dots, x_n\}$. Now, by Theorem 5.3.9, there exist affinely independent $y_0, \dots, y_n \in \mathbb{R}^n$ such that $\|x_i - x_j\| = \|y_i - y_j\|^2$. By letting $r > 0$ be the circumradius of the y_i , it follows from Lemma 6.1.4 that for all $w_0, \dots, w_n \in \mathbb{R}$ such that $\sum_{i=0}^n w_i = 1$,

$$\sum_{i=0}^n \sum_{j=0}^n w_i w_j d(x_i, x_j) = \sum_{i=0}^n \sum_{j=0}^n w_i w_j \|y_i - y_j\|^2 = 2r^2 - 2 \left\| \sum_{i=0}^n w_i y_i \right\|^2,$$

and $\sum_{i=0}^n w_i y_i = 0$ for a unique choice of the w_i . Hence $\overline{M}(X, d) = 2r^2$ and the signed measure $\mu = \sum_{i=0}^n w_i \delta_{x_i} \in \mathcal{M}(X)$ is \overline{M} -maximal when the w_i are the unique barycentric coordinates such that $\sum_{i=0}^n w_i y_i = 0$. Given that all signed measures on X are atomic, it follows that an \overline{M} -maximal measure on X must be unique.

Let $s = \inf \left\| \sum_{i=0}^n w_i y_i \right\| \geq 0$, where the infimum is taken over all $w_0, \dots, w_n \geq 0$ such that $\sum_{i=0}^n w_i = 1$. Then $M(X) = 2(r^2 - s^2)$. Now, assume without loss of generality that the y_i are chosen such that their circumcentre is the origin. As the convex hull of the y_i consists precisely of the points $\sum_{i=0}^n v_i y_i$ such that $v_0, \dots, v_n \geq 0$ and $\sum_{i=0}^n v_i = 1$, the distance from the origin to this convex hull is then $\inf \left\| \sum_{i=0}^n v_i y_i - 0 \right\| = s$. We already know from Theorem 6.2.1 that a unique M -maximal measure exists. □

We now use work of Gower to calculate the circumradius of a $1/2$ -embedding of a strictly quasihypermetric space X in a Euclidean space of minimal dimension. This result will give a formula for $\overline{M}(X)$ which may be calculated on a computer.

The following result is due to Gower [23, pg. 82-83].

Lemma 6.3.2. *Let $n \in \mathbb{N}$, let A be a real symmetric $n \times n$ matrix with diagonal entries zero, let $e \in \mathbb{R}^n$ be the vector consisting entirely of ones, let $s \in \mathbb{R}^n$ such that $\langle s, e \rangle = 1$ and let F be the $n \times n$ matrix*

$$F = -(I - es^T)A(I - se^T).$$

Then A is almost-negative definite if and only if F is positive semi-definite.

Proof. Let $t \in \mathbb{R}^n$ such that $\langle t, e \rangle = 1$ and let G be the $n \times n$ matrix

$$G = -(I - et^T)A(I - te^T).$$

Then as $\langle s, e \rangle = \langle e, s \rangle = s^T e = e^T s = 1$, we have that

$$\begin{aligned} (I - es^T)(I - et^T) &= I - et^T - es^T + es^T et^T = I - es^T, \\ (I - te^T)(I - se^T) &= I - se^T - te^T + te^T se^T = I - se^T. \end{aligned}$$

It follows that for all $x \in \mathbb{R}^n$,

$$\begin{aligned} x^T F x &= -x^T (I - es^T) A (I - se^T) x \\ &= -x^T (I - es^T) (I - et^T) A (I - te^T) (I - se^T) x \\ &= x^T (I - es^T) G (I - se^T) x \\ &= ((I - se^T)x)^T G (I - se^T)x. \end{aligned}$$

Note that $(I - se^T)x \in \mathbb{R}^n$. Therefore, if G is positive semi-definite then F is positive semi-definite. The converse of this statement is shown similarly.

Suppose that $t = (0, \dots, 0, 1)$, which is consistent with the requirement that $\langle t, e \rangle = 1$. Then $G_{ij} = -(A_{ij} - A_{in} - A_{nj})$ and by Lemma 5.3.7, G is positive semi-definite if and only if A is almost-negative definite. The result follows. \square

The proof of the following result is given by Baxter [7, pg. 3].

Lemma 6.3.3. *Let $n \in \mathbb{N}$ and let A be a $n \times n$ strictly almost-negative definite matrix with diagonal entries zero. Then A is non-singular.*

Proof. Seeing as a 1×1 matrix is not strictly almost-negative definite, assume that $n \geq 2$. As A is a symmetric matrix which gives a negative definite quadratic form on a subspace of dimension $n - 1$, as in the proof of Lemma 5.3.6 it can be shown that A has at least $n - 1$ negative eigenvalues. Given that the trace of A is zero, the sum of the eigenvalues of A is zero, and A must have a positive eigenvalue. We have found a complete set of eigenvalues for A , whose product is $\det A$. Hence the sign of $\det A$ is given by the sign of $(-1)^{n-1} \neq 0$, giving that A is non-singular. \square

The following result uses a proof essentially due to Gower [23, pg. 83-84]. (Gower was concerned with finding the circumradius of a finite set of points in a Euclidean space.)

Theorem 6.3.4. *Let (X, d) be a finite strictly quasihypermetric space with $n \geq 2$ points, and let D be a distance matrix of X . Then D is non-singular and*

$$\overline{M}(X) = \left(\sum_{i=1}^n \sum_{j=1}^n (D^{-1})_{ij} \right)^{-1}.$$

Proof. Write X as $X = \{x_1, \dots, x_n\}$ and let D be the $n \times n$ matrix such that $D_{ij} = d(x_i, x_j)$. Then D is strictly almost-negative definite, and the previous result gives that D is non-singular. Let $e \in \mathbb{R}^n$ be the vector consisting entirely of ones, let $s \in \mathbb{R}^n$ such that $\langle s, e \rangle = 1$ and let

$$F = -(I - es^T)D(I - se^T).$$

By Lemma 6.3.2, F is positive semi-definite and by Lemma 5.3.6 there exists an $n \times n$ matrix Y such that $F = YY^T$. Let $y_1, \dots, y_n \in \mathbb{R}^n$ be such that $\sqrt{2}y_1, \dots, \sqrt{2}y_n$ are the rows of Y . Now, the columns of the matrix se^T are s and the rows of the matrix es^T are s^T . Using

$$YY^T = -(I - es^T)D(I - se^T) = -D + Dse^T + es^TD - es^TDse^T,$$

we have that for all i and j ,

$$\begin{aligned} 2 \langle y_i, y_j \rangle &= -D_{ij} + \sum_{k=1}^n s_k D_{ik} + \sum_{k=1}^n s_k D_{kj} - s^T D s \\ &= -D_{ij} + \sum_{k=1}^n s_k (D_{ik} + D_{kj}) - s^T D s. \end{aligned}$$

Hence for all i and j , $\|y_i\|^2 = \sum_{k=1}^n s_k D_{ik} - (1/2)s^T D s$ and

$$\begin{aligned} \|y_i - y_j\|^2 &= \|y_i\|^2 + \|y_j\|^2 - 2 \langle y_i, y_j \rangle \\ &= \sum_{k=1}^n s_k (D_{ik} + D_{jk}) - s^T D s + D_{ij} - \sum_{k=1}^n s_k (D_{ik} + D_{kj}) + s^T D s \\ &= D_{ij}. \end{aligned}$$

Therefore the y_i form a $1/2$ -embedding of the x_i in \mathbb{R}^n . By applying a suitable translation followed by a rotation, there exist $y'_1, \dots, y'_n \in \mathbb{R}^{n-1}$ such that the y_i are isometric to the y'_i . Hence the y'_i are a $1/2$ -embedding of the x_i in \mathbb{R}^{n-1} , and the y'_i must lie on a sphere of some radius $r > 0$, giving that $\overline{M}(X) = 2r^2$. Given that the image of a sphere under an isometry is a sphere of the same radius, the y_i must also lie on a sphere of radius r .

Note that F and Y are dependent on s , but r is not. We now want to show that s can be chosen such that the y_i lie on a sphere centred at the origin. Let $(z_1, \dots, z_n)^T = 2Ds - (s^T D s)e \in \mathbb{R}^n$, which is the diagonal of F . Then for each i ,

$$z_i = 2\|y_i\|^2 = 2 \sum_{k=1}^n s_k D_{ik} - s^T D s,$$

and so a sphere containing the y_i of radius r will be centred at the origin if and only if $2Ds - (s^T D s)e = 2r^2 e$. Given that $\det D \neq 0$ and that D^{-1} is symmetric, this equation holds if and only if

$$s = \frac{(2r^2 + s^T D s)D^{-1}e}{2}.$$

Given that $(2r^2 + s^T D s)/2 \in \mathbb{R}$, we look for a solution of the form $s = kD^{-1}e$ for some $k \in \mathbb{R}$. Given that $s = 0$ is not a solution to our equation, assume that $k \neq 0$.

Using $\langle s, e \rangle = \langle e, s \rangle = 1$, it follows that

$$\frac{e^T D^{-1}e}{k} = 1,$$

and so $2Ds - (s^T Ds)e = 2r^2 e$ if and only if

$$s = \frac{D^{-1}e}{e^T D^{-1}e},$$

noting that $e^T D^{-1}e \neq 0$. By substituting, we then obtain

$$2D \frac{D^{-1}e}{e^T D^{-1}e} - \left(\frac{e^T D^{-1}}{e^T D^{-1}e} D \frac{D^{-1}e}{e^T D^{-1}e} \right) e = 2r^2 e,$$

which gives that

$$r^2 = \frac{1}{2e^T D^{-1}e} = \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n (D^{-1})_{ij} \right)^{-1}.$$

Finally, we have

$$\overline{M}(X) = \left(\sum_{i=1}^n \sum_{j=1}^n (D^{-1})_{ij} \right)^{-1}. \quad \square$$

Using the fact that there exists an algorithm for calculating the inverse of an $n \times n$ non-singular matrix in running time of order $O(n^3)$, it can be seen that for a finite metric space X with n points, $\overline{M}(X)$ can be calculated on a computer in running time of order $O(n^3)$.

Finally, we extend our result to certain infinite strictly quasihypermetric spaces.

Theorem 6.3.5. *Let (X, d) be a compact strictly quasihypermetric space. Then $\overline{M}(X)$ is finite if and only if each finite subset A of X is 1/2-embeddable on a sphere in ℓ_2 of minimal radius $r(A)$ and $r = \sup r(A) < \infty$. Further, if $\overline{M}(X)$ is finite then*

$$M(X) = 2(r^2 - s^2) \quad \text{and} \quad \overline{M}(X) = 2r^2,$$

where s is the distance from the centre of a sphere in ℓ_2 of radius r containing a 1/2-embedding of X to the convex hull of the 1/2-embedding.

Proof. Let $n \in \mathbb{N}_0$ and let $x_0, \dots, x_n \in X$. Then there exists an affinely independent 1/2-embedding $y_0, \dots, y_n \in \mathbb{R}^n$ of the x_i in \mathbb{R}^n . Let $r(x_0, \dots, x_n)$ be the circumradius of y_0, \dots, y_n and let $s(x_0, \dots, x_n)$ be the distance from the circumcentre of y_0, \dots, y_n to the convex hull of these points. Then by Theorem 6.3.1,

$$\begin{aligned} M(\{x_0, \dots, x_n\}) &= 2(r(x_0, \dots, x_n)^2 - s(x_0, \dots, x_n)^2), \\ \overline{M}(\{x_0, \dots, x_n\}) &= 2r(x_0, \dots, x_n)^2. \end{aligned}$$

Let $R \times S$ be the collection of all such $(r(x_0, \dots, x_n), s(x_0, \dots, x_n))$ as n ranges over \mathbb{N} and x_0, \dots, x_n range over X . Given that

$$M(X) = \sup_{A \in \mathcal{P}'(X)} M(A) \quad \text{and} \quad \overline{M}(X) = \sup_{A \in \mathcal{P}'(X)} \overline{M}(A),$$

where $\mathcal{P}'(X)$ denotes the family of finite subsets of X , we then have

$$M(X) = \sup_{(r,s) \in R \times S} 2(r^2 - s^2) \quad \text{and} \quad \overline{M}(X) = \sup_{r \in R} 2r^2.$$

It follows that $\overline{M}(X) < \infty$ if and only if R is bounded, and by Theorem 6.3.1, R is bounded if and only if X is $1/2$ -embeddable on a sphere in ℓ_2 .

Suppose that $\overline{M}(X)$ is finite, and let $f: X \rightarrow \ell_2$ be a function such that $f(X)$ is a $1/2$ -embedding of X on a sphere in ℓ_2 . Then R is bounded, and by letting $r = \sup R < \infty$, we have that r is the circumradius of $f(X)$, and so Theorem 6.3.1 gives that $\overline{M}(X) = 2r^2$. Let c be the circumcentre of $f(X)$, and let

$$s = \inf \left\| c - \sum_{i=1}^n w_i f(x_i) \right\|$$

be the distance from c to the convex hull of $f(X)$, where the infimum is taken over all $n \in \mathbb{N}$ and all $w_1, \dots, w_n \geq 0$ such that $\sum_{i=1}^n w_i = 1$ and all $x_1, \dots, x_n \in X$. We want to show that $M(X) = 2(r^2 - s^2)$.

Using Lemma 2.3.23 and a recursive construction, for each $n \in \mathbb{N}$, let $m(n) \in \mathbb{N}$ and let $A_{1,n}, \dots, A_{m(n),n} \in \mathcal{B}(X)$ be a partition of X such that $D(A_{i,n}) < 1/n$ for each $i = 1, \dots, m(n)$ and each successive partition is a refinement of the previous partition. Now for each $n \in \mathbb{N}$, let $X_n = \{x_{1,n}, \dots, x_{m(n),n}\} \subseteq X$ be such that $x_{i,n} \in A_{i,n}$ for each $i = 1, \dots, m(n)$. By properties of the $A_{i,j}$, we may choose the $x_{i,j}$ such that $X_n \subseteq X_{n+1}$ for each $n \in \mathbb{N}$. Note that it can easily be shown that $\bigcup_{n=1}^{\infty} X_n$ is dense in X . Hence $X_n \rightarrow X$ with respect to the Hausdorff metric on $H(X)$, and by Theorem 4.3.7 it follows that $M(X_n) \rightarrow M(X)$.

We now show that $\overline{M}(X_n) \rightarrow \overline{M}(X)$. Let $\varepsilon > 0$ and let $(Y_n) \in \mathcal{P}'(X)$ be a sequence of finite subsets of X chosen such that $(\overline{M}(Y_n))$ is non-decreasing and $\overline{M}(Y_n) \rightarrow \overline{M}(X)$. Then there exists $n_0 \in \mathbb{N}$ such that $|\overline{M}(Y_{n_0}) - \overline{M}(X)| < \varepsilon/3$.

Writing $Y_{n_0} = \{y_0, \dots, y_m\}$, by Theorem 6.3.1 there exists an \overline{M} -maximal measure $\sum_{i=1}^m v_i \delta_{y_i}$ on Y_{n_0} . Let $v = \max \{|v_i| : i = 1, \dots, m\} > 0$, let $x_0, \dots, x_m \in \bigcup_{n=1}^\infty X_n$ be such that for each $i = 1, \dots, m$, $d(y_i, x_i) < \varepsilon/3m^2v^2$, and let $n_1 \in \mathbb{N}$ be such that $x_0, \dots, x_m \in X_{n_1}$. We then have that $\sum_{i=1}^m v_i \delta_{x_i} \in \mathcal{M}(X_{n_1})$ is a signed measure of unit mass, and as

$$\begin{aligned}
& \left| I \left(\sum_{i=1}^m v_i \delta_{x_i} \right) - \overline{M}(X) \right| \\
& \leq \left| I \left(\sum_{i=1}^m v_i \delta_{x_i} \right) - \overline{M}(Y_{n_0}) \right| + |\overline{M}(Y_{n_0}) - \overline{M}(X)| \\
& < \left| \sum_{i=1}^m \sum_{j=1}^m v_i v_j d(x_i, x_j) - \sum_{i=1}^m \sum_{j=1}^m v_i v_j d(y_i, y_j) \right| + \frac{\varepsilon}{3} \\
& \leq \sum_{i=1}^m \sum_{j=1}^m v^2 (|d(x_i, x_j) - d(x_i, y_j)| + |d(x_i, y_j) - d(y_i, y_j)|) + \frac{\varepsilon}{3} \\
& \leq \sum_{i=1}^m \sum_{j=1}^m v^2 (d(x_j, y_j) + d(x_i, y_i)) + \frac{\varepsilon}{3} \\
& < \sum_{i=1}^m \sum_{j=1}^m v^2 \left(\frac{\varepsilon}{3m^2v^2} + \frac{\varepsilon}{3m^2v^2} \right) + \frac{\varepsilon}{3} \\
& = \varepsilon,
\end{aligned}$$

it follows that

$$\overline{M}(X) - \varepsilon < I \left(\sum_{i=1}^m v_i \delta_{x_i} \right) \leq \overline{M}(X_{n_1}) \leq \overline{M}(X).$$

Since $(\overline{M}(X_n))$ is non-decreasing, it must then be the case that $\overline{M}(X_n) \rightarrow \overline{M}(X)$.

Let $(c_n) \in \ell_2$ and $(r_n) \in \mathbb{R}$ and $(s_n) \in \mathbb{R}$ be sequences such that for each $n \in \mathbb{N}$, the circumsphere of $f(X_n)$ is of minimal radius r_n and its circumcentre c_n is distance s_n from the convex hull of $f(X_n)$. As in the proof of Theorem 5.3.15, we may assume that f has been constructed such that $c_n \rightarrow c$. We then have that

$$\begin{aligned}
\lim_{n \rightarrow \infty} M(X_n) &= \lim_{n \rightarrow \infty} 2(r_n^2 - s_n^2) = M(X), \\
\lim_{n \rightarrow \infty} \overline{M}(X_n) &= \lim_{n \rightarrow \infty} 2r_n^2 = 2r^2 = \overline{M}(X),
\end{aligned}$$

from which we can deduce that $r_n \rightarrow r$ and (s_n) is convergent. Letting $t \in \mathbb{R}$ be the limit of (s_n) , it remains to be shown that $t = s$.

Firstly, we show that $s \leq t$. For each $n \in \mathbb{N}$, the convex hull C_n of $f(X_n)$ is compact, and so there exists $\sigma_n \in C_n$ such that $s_n = \|c_n - \sigma_n\|$. Let $\varepsilon > 0$ and let $n_0 \in \mathbb{N}$ be such that for all $n \geq n_0$, $\|c - c_n\| < \varepsilon/2$ and $s_n = \|c_n - \sigma_n\| < t + \varepsilon/2$. Then for all $n \geq n_0$, given that C_n is contained in the convex hull of $f(X)$, by the definition of s we must have that

$$s \leq \|c - \sigma_n\| \leq \|c - c_n\| + \|c_n - \sigma_n\| < \frac{\varepsilon}{2} + t + \frac{\varepsilon}{2} = t + \varepsilon,$$

from which it follows that $s \leq t$.

To complete the proof, it remains to be shown that $t \geq s$. Let $\varepsilon > 0$, let C be the convex hull of $f(X)$, and let $\sigma \in C$ such that $|\|c - \sigma\| - s| < \varepsilon/3$. Since it is obvious that f is continuous, $f(\bigcup_{n=1}^{\infty} X_n)$ is dense in $f(X)$. Writing $\sigma = \sum_{i=1}^n w_i f(x_i)$ for some $n \in \mathbb{N}$ and some $w_1, \dots, w_n > 0$ such that $\sum_{i=1}^n w_i = 1$ and some $x_1, \dots, x_n \in X$, let $\sigma' = \sum_{i=1}^n w_i f(y_i) \in C$, where each $y_i \in \bigcup_{n=1}^{\infty} X_n$ is chosen such that $\|f(x_i) - f(y_i)\| \leq \varepsilon/3$. Then

$$\begin{aligned} \|\sigma - \sigma'\| &= \left\| \sum_{i=1}^n w_i f(x_i) - \sum_{i=1}^n w_i f(y_i) \right\| \\ &\leq \sum_{i=1}^n w_i \|f(x_i) - f(y_i)\| \\ &< \sum_{i=1}^n w_i \frac{\varepsilon}{3} \\ &= \frac{\varepsilon}{3}. \end{aligned}$$

Now, let $m_0 \in \mathbb{N}$ such that $\sigma' \in C_{m_0}$, and let $m_1 \in \mathbb{N}$ such that $\|c_n - c\| < \varepsilon/3$ whenever $n \geq m_1$. Then for all $m \in \mathbb{N}$ such that $m \geq \max\{m_0, m_1\}$, $\sigma' \in C_m$, and so

$$s_n \leq \|c_m - \sigma'\| \leq \|c_m - c\| + \|c - \sigma\| + \|\sigma - \sigma'\| < \frac{\varepsilon}{3} + s + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = s + \varepsilon.$$

It is then a routine exercise to show that $t \leq s$, as t is the limit of (s_n) . \square

Alexander and Stolarsky [2, pg. 14] used this result to calculate $\overline{M}([a, b])$. Note that we calculated this average distance from first principles in Proposition 1.2.10.

Proposition 6.3.6. *Let $a, b \in \mathbb{R}$ such that $a < b$. Then $\overline{M}([a, b]) = (b - a)/2$.*

Proof. It will be sufficient to prove the result for $a = 0$ and $b = 1$. By a simple consideration of the triangle inequality, any sphere in ℓ_2 containing a $1/2$ -embedding of $[0, 1]$ must be of radius at least $1/2$. It will be sufficient to show that each finite collection of points in $[0, 1]$ can be $1/2$ -embedded on a sphere in ℓ_2 of radius at most $1/2$. Let $x_1, \dots, x_n \in [0, 1]$ be ordered and for each i , let

$$y_i = (\sqrt{x_1}, \dots, \sqrt{x_i - x_{i-1}}, 0, \dots) \in \ell_2.$$

Then for each i and j such that $i > j$,

$$\begin{aligned} \|y_i - y_j\| &= \|(0, \dots, 0, \sqrt{x_{j+1} - x_j}, \dots, \sqrt{x_i - x_{i-1}}, 0, \dots)\| \\ &= \left(\sum_{k=j+1}^i (x_k - x_{k-1}) \right)^{1/2} \\ &= |x_i - x_j|^{1/2}. \end{aligned}$$

Hence the y_i form a $1/2$ -embedding of the x_i in ℓ_2 . Now, let

$$y = \frac{1}{2} (\sqrt{x_1}, \sqrt{x_2 - x_1}, \dots, \sqrt{x_n - x_{n-1}}, 0, \dots) \in \ell_2.$$

Then for each i ,

$$\begin{aligned} \|y_i - y\| &= \left\| \frac{1}{2} (\sqrt{x_1}, \dots, \sqrt{x_i - x_{i-1}}, -\sqrt{x_i - x_{i-1}}, \dots, -\sqrt{x_n - x_{n-1}}, 0, \dots) \right\| \\ &= \frac{1}{2} \left(x_1 + \sum_{k=2}^n (x_k - x_{k-1}) \right)^{1/2} \\ &= \frac{1}{2} x_n^{1/2}. \end{aligned}$$

Hence the y_i lie on a sphere centred at y of radius $(1/2)x_n^{1/2} \leq 1/2$. □

6.4 Some values of $m(X)$, $M(X)$ and $\overline{M}(X)$

We conclude this chapter by showing some of our previous calculations of Gross-Stadje numbers also give the values of $M(X)$ and $\overline{M}(X)$ for various concrete spaces X . We also discuss max-min, M -maximal and \overline{M} -maximal measures on regular polygons.

It is immediate from Proposition 4.5.7 and Theorem 6.2.10 that:

Proposition 6.4.1. *Let $n \in \mathbb{N}$ such that $n \geq 2$, let d be the usual metric on S^{n-1} and let μ be the usual normalised measure on S^{n-1} . Then μ is a unique d -invariant measure on S^{n-1} and*

$$m(S^{n-1}) = M(S^{n-1}) = \overline{M}(S^{n-1}) = \frac{2^{n-1} \left(\Gamma\left(\frac{n}{2}\right)\right)^2}{\sqrt{\pi} \Gamma\left(\frac{2n-1}{2}\right)}.$$

where Γ denotes the gamma function.

We can calculate $m(S^{n-1})$ for $n = 2, \dots, 6$ as

n	2	3	4	5	6
$m(S^{n-1})$	$2/\pi$	$4/3$	$64/15\pi$	$48/35$	$4096/945\pi$

Proposition 4.5.8 calculated the Gross-Stadje number of compact connected arcs of S^1 . It is immediate from this Proposition and Theorem 6.2.10 that:

Proposition 6.4.2. *Let X_ϕ be an arc of S^1 subtending an angle $0 < \phi \leq 2\pi$ at the centre, let d be the usual metric on X_ϕ and let μ be the usual normalised measure on X_ϕ . Then μ is the unique d -invariant measure on X_ϕ and*

$$m(X_\phi) = M(X_\phi) = \overline{M}(X_\phi) = \frac{8}{\phi + \cot(\phi/4)}.$$

Recall that Proposition 4.5.4 gives that the Borel probability measure uniformly distributed on the vertices of a polygon is a max-min measure. We now show that this measure is also the unique M -maximal measure.

Proposition 6.4.3. *Let (X, d) be a regular polygon. Then the probability measure uniformly distributed on the vertices of a regular polygon is both a max-min and the unique M -maximal measure on X . Further, $m(X) < M(X) < \overline{M}(X)$.*

Proof. Let $n \in \mathbb{N}$ such that $n \geq 3$, let X be a regular n -gon and let x_0, \dots, x_{n-1} be the vertices of X . It follows from Corollary 4.5.14 that there exists an M -maximal measure $\mu_0 \in \mathcal{M}^1(X)$ supported on the x_i . Hence, there exist $w_0, \dots, w_{n-1} \geq 0$ such that $\mu = \sum_{i=0}^{n-1} w_i \delta_{x_i}$ and $\sum_{i=0}^{n-1} w_i = 1$. We have that

$$M(X) = I(\mu) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} w_i w_j \|x_i - x_j\|.$$

Now, for each $i = 1, \dots, n-1$, let $\mu_i \in \mathcal{M}^1(X)$ be the measure

$$\mu_i = \sum_{j=0}^{n-1} w_{i+j} \delta_{x_j},$$

where addition of the weight indexes is addition modulo n , and let

$$\mu = (1/n) \sum_{i=0}^{n-1} \mu_i \in \mathcal{M}^1(X).$$

Noting that X is strictly quasihypermetric, repeated applications of Theorem 5.4.5 give that $I(\mu_0) \leq I(\mu)$ with equality occurring if and only if $\mu_0 = \mu_i$ for each $i = 1, \dots, n-1$ if and only if $w_i = 1/n$ for each $i = 0, \dots, n-1$. It follows that the measure uniformly distributed on the x_i is the M -maximal measure.

That $m(X) < M(X) < \overline{M}(X)$ follows from Theorem 6.2.10 after noting that μ is not d -invariant. \square

As a corollary to the proof of the previous result, we obtain:

Corollary 6.4.4. *Let (X, d) be a finite metric space consisting of points evenly spaced around a circle. Then the probability measure uniformly distributed on the points of X is d -invariant and hence the unique M -maximal measure on X and the unique \overline{M} -maximal measure on X . Further, $M(X) = \overline{M}(X)$.*

6.5 Summary

Chapter 4 discussed certain properties of $m(X)$, $M(X)$ and $\overline{M}(X)$, and this chapter has continued this investigation for the case of a compact quasihypermetric space X . We saw that the quasihypermetric property is necessary to obtain $\overline{M}(X) < \infty$, but that the converse does not hold. It was seen that $\overline{M}(X) < \infty$ when X is finite and strictly quasihypermetric or when X is Euclidean.

For compact quasihypermetric spaces (X, d) we investigated the significance of d -invariant measures on X and their relationship to demonstrating existence of M -maximal measures on X and the existence of \overline{M} -maximal signed measures on X .

When X is also connected, the relationship between $m(X)$, $M(X)$ and $\overline{M}(X)$ is related to the existence of a d -invariant probability measure on X , and we saw that all three averages are equal or none are equal.

Geometrical interpretations of $M(X)$ and $\overline{M}(X)$ were investigated in the case of X being compact quasihypermetric, and certain formulas for them were established in terms of the circumradius of a $1/2$ -embedding of X in ℓ_2 and the distance from the convex hull of the $1/2$ -embedding to its circumcentre.

Finally, we extended results from Chapter 4 to calculate $m(X)$, $M(X)$ and $\overline{M}(X)$ for all spheres X and to calculate the M -maximal measure for regular polygons.

Chapter 7

Other work

We conclude this thesis by noting other work which we may have chosen to consider in our investigation of average distances in compact metric spaces. Given a compact (connected) metric space X , simple considerations of $m(X)$, $M(X)$ and $\overline{M}(X)$ quickly lead to many different questions which may be discussed. This thesis has only discussed a small subset of the questions which immediately arise, and would rapidly expand into several volumes if we attempted to consider many more of them.

7.1 Averaging powers of distances

Recall that the most general form of the Gross-Stadje Theorem states that for all compact connected Hausdorff spaces X and all continuous symmetric functions $f: X \times X \rightarrow \mathbb{R}$, there exists a unique $m(X, f) \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ and for all $x_1, \dots, x_n \in X$ there exists $y \in X$ such that

$$\sum_{i=1}^n f(x_i, y) = m(X, f) .$$

As a consequence, we could have chosen to investigate the averages $m(X, d^\lambda)$, $M(X, d^\lambda)$ and $\overline{M}(X, d^\lambda)$ of the continuous symmetric function $d^\lambda: X \times X \rightarrow \mathbb{R}$ in a compact (connected) metric space (X, d) for any $\lambda > 0$. Björck [8, pg. 258] discusses $M(X, d^\lambda)$ for $0 < \lambda < 2$, and Alexander and Stolarsky [2] discuss $\overline{M}(X, d^\lambda)$ for Euclidean spaces X and $0 < \lambda \leq 1$.

It has been noted by Cleary et al. [16, pg. 264] that for compact connected subsets (X, d) of Euclidean spaces, it is easier to calculate $m(X, d^2)$ than $m(X, d)$, given that the metric on such spaces arises from an inner-product. It is reasonable to expect that this would also be the case for $M(X, d^2)$ and $\overline{M}(X, d^2)$.

Wilson [54] has calculated $m(X, d^2)$ for spaces such as the ellipse or the Reuleaux triangle, where $m(X, d)$ is not known, and for regular polygons.

7.2 Bounds for $m(X)$, $M(X)$ and $\overline{M}(X)$

Yost [57] discusses bounds for $m(X)$ where X is a compact connected subset of a finite dimensional normed space. In particular, he shows that when V is any finite dimensional normed space then there exists a constant $k(V) < 1$ such that $m(X) \leq k(V)$ whenever X is a compact connected subset of V of unit diameter, and that $\{m(X) : X \text{ is a compact connected subset of } V \text{ such that } D(X) = 1\}$ is the interval $[1/2, k(V)]$. It would be interesting to determine if a corresponding property holds for the average distances $M(X)$ and $\overline{M}(X)$.

An application of Jung's Theorem gives that a compact metric space contained in some \mathbb{R}^n can be embedded in the convex hull of a sphere of diameter $\sqrt{2n/(n+1)}$. Nickolas and Yost [41] combined this with the known value of $m(S^{n-1})$ to give the bounds $n/(n+1) \leq k(\mathbb{R}^n) < \sqrt{n/(n+1)}$ for each $n \in \mathbb{N}$. For $n = 2$, using a similar technique they obtain the improved estimate

$$0.6675276 < k(\mathbb{R}^2) < \frac{2 + \sqrt{3}}{3\sqrt{3}} < 0.7182336.$$

It is noted by Yost [57] that the above lower bound for $k(\mathbb{R}^2)$ is a lower bound for $m(X)$, where X is the Reuleaux triangle of unit diameter. It would be of interest to know the Gross-Stadje number of this space, and if it does give rise to $k(\mathbb{R}^2)$.

7.3 Examples

This thesis mainly studied measure theoretic and analytic questions which arise from considering $m(X)$, $M(X)$ and $\overline{M}(X)$ and did not include many calculations of these average distances for concrete spaces. Exact values of these averages are only known for a small collection of spaces. The value of $\overline{M}(X)$ is not even known for a space as simple as an equilateral triangle, nor is $M(X)$ known for an arbitrary triangle! It would be worth carrying out further work calculating exact values of $m(X)$, $M(X)$ and $\overline{M}(X)$ for some well-known spaces.

Bibliography

- [1] Ralph Alexander. Two notes on metric geometry. *Proc. Amer. Math. Soc.*, 64(2):317–320, 1977.
- [2] Ralph Alexander and Kenneth B. Stolarsky. Extremal problems of distance geometry related to energy integrals. *Trans. Amer. Math. Soc.*, 193:1–31, 1974.
- [3] Howard Anton. *Calculus with analytic geometry*. John Wiley & Sons Inc., New York, 1995.
- [4] E. Artin. *Geometric algebra*. Interscience Publishers, Inc., New York-London, 1957.
- [5] Patrice Assouad. Un espace hypermétrique non plongeable dans un espace L_1 . *C. R. Acad. Sci. Paris Sér. A-B*, 285(5):A361–A363, 1977.
- [6] Robert G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1995.
- [7] B. J. C. Baxter. Conditionally positive functions and p -norm distance matrices. *Constr. Approx.*, 7(4):427–440, 1991.
- [8] Göran Björck. Distributions of positive mass, which maximize a certain generalized energy integral. *Ark. Mat.*, 3:255–269, 1956.
- [9] Leonard M. Blumenthal. *Theory and applications of distance geometry*. Second edition. Chelsea Publishing Co., New York, 1970.

- [10] N. Bourbaki. *Éléments de mathématique. Fasc. XIII. Livre VI: Intégration. Chapitres 1, 2, 3 et 4.* Deuxième édition revue et augmentée. Actualités Scientifiques et Industrielles, No. 1175. Hermann, Paris, 1965.
- [11] Jean Bretagnolle, Didier Dacunha-Castelle, and Jean-Louis Krivine. Lois stables et espaces L^p . *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 2:231–259, 1965/1966.
- [12] Arne Brøndsted. *An introduction to convex polytopes*, volume 90 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1983.
- [13] Gustave Choquet. *Lectures on analysis. Vol. I: Integration and topological vector spaces.* Edited by J. Marsden, T. Lance and S. Gelbart. W. A. Benjamin, Inc., New York-Amsterdam, 1969.
- [14] Gustave Choquet. *Lectures on analysis. Vol. II: Representation theory.* Edited by J. Marsden, T. Lance and S. Gelbart. W. A. Benjamin, Inc., New York-Amsterdam, 1969.
- [15] Joan Cleary. Numerical geometry, honours project. *La Trobe University Pure Mathematics Research Paper*, 83-2, 1982.
- [16] Joan Cleary, Sidney A. Morris, and David Yost. Numerical geometry—numbers for shapes. *Amer. Math. Monthly*, 93(4):260–275, 1986.
- [17] Donald L. Cohn. *Measure theory*. Birkhäuser Boston Inc., Boston, MA, 1993.
- [18] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups*, volume 290 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, third edition, 1999. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.
- [19] E. T. Copson. *Metric spaces*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 57. Cambridge University Press, London, 1968.

- [20] H. S. M. Coxeter. *Introduction to geometry*. John Wiley & Sons Inc., New York, 1969.
- [21] H. G. Eggleston. *Convexity*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 47. Cambridge University Press, New York, 1958.
- [22] Ryszard Engelking. *General topology*, volume 6 of *Sigma Series in Pure Mathematics*. Heldermann Verlag, Berlin, 1989.
- [23] J. C. Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Appl.*, 67:81–97, 1985.
- [24] O. Gross. The rendezvous value of a metric space. In *Advances in game theory*, pages 49–53. Princeton Univ. Press, Princeton, N.J., 1964.
- [25] Felix Hausdorff. *Set theory*. Second edition. Translated from the German by John R. Aumann et al. Chelsea Publishing Co., New York, 1962.
- [26] John B. Kelly. Metric inequalities and symmetric differences. In *Inequalities, II (Proc. Second Sympos., U.S. Air Force Acad., Colo., 1967)*, pages 193–212. Academic Press, New York, 1970.
- [27] John B. Kelly. Hypermetric spaces and metric transforms. In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, pages 149–158. Academic Press, New York, 1972.
- [28] John B. Kelly. Hypermetric spaces. In *The geometry of metric and linear spaces (Proc. Conf., Michigan State Univ., East Lansing, Mich., 1974)*, pages 17–31. Lecture Notes in Math., Vol. 490. Springer, Berlin, 1975.
- [29] Paul J. Kelly and Max L. Weiss. *Geometry and convexity*. John Wiley & Sons, New York-Chichester-Brisbane, 1979.
- [30] André E. Kézdy, Mark J. Nielsen, and Hunter S. Snevily. Generalized triangle inequalities in \mathbb{R}^n . *Bull. Inst. Combin. Appl.*, 33:23–28, 2001.

- [31] Alexander Koldobsky and Yossi Lonke. A short proof of Schoenberg's conjecture on positive definite functions. *Bull. London Math. Soc.*, 31(6):693–699, 1999.
- [32] K. Kuratowski. *Topology. Vol. I*. New edition, revised and augmented. Translated from the French by J. Jaworowski. Academic Press, New York, 1966.
- [33] G. Larcher, W. Ch. Schmid, and R. Wolf. On the approximation of certain mass distributions appearing in distance geometry. *Acta Math. Hungar.*, 87(4):295–316, 2000.
- [34] Steven R. Lay. *Convex sets and their applications*. John Wiley & Sons Inc., New York, 1982.
- [35] L. Lovász, L. Pyber, D. J. A. Welsh, and G. M. Ziegler. Combinatorics in pure mathematics. In *Handbook of combinatorics, Vol. 1, 2*, pages 2039–2082. Elsevier, Amsterdam, 1995.
- [36] L. Mirsky. *An introduction to linear algebra*. Dover Publications Inc., New York, 1990.
- [37] Sidney A. Morris. *Topology without tears*. Unpublished, 2004.
- [38] Sidney A. Morris and Peter Nickolas. On the average distance property of compact connected metric spaces. *Arch. Math. (Basel)*, 40(5):459–463, 1983.
- [39] Peter M. Neumann, Gabrielle A. Stoy, and Edward C. Thompson. *Groups and geometry*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1994.
- [40] Peter Nickolas and Reinhard Wolf. Distance geometry, the quasihypermetric property and spaces of measures. *In preparation*.
- [41] Peter Nickolas and David Yost. The average distance property for subsets of Euclidean space. *Arch. Math. (Basel)*, 50(4):380–384, 1988.

- [42] Hukukane Nikaidô. On von Neumann's minimax theorem. *Pacific J. Math.*, 4:65–72, 1954.
- [43] K. R. Parthasarathy. *Probability measures on metric spaces*. Probability and Mathematical Statistics, No. 3. Academic Press Inc., New York, 1967.
- [44] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev. *Integraly i ryady. Tom 1*. Fiziko-Matematicheskaya Literatura, Moscow, 2003.
- [45] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, 1966.
- [46] I. J. Schoenberg. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert" [MR1503246]. *Ann. of Math. (2)*, 36(3):724–732, 1935.
- [47] I. J. Schoenberg. On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space. *Ann. of Math. (2)*, 38(4):787–793, 1937.
- [48] Wolfgang Stadge. A property of compact connected spaces. *Arch. Math. (Basel)*, 36(3):275–280, 1981.
- [49] Frank M. Stewart. *Introduction to linear algebra*. Van Nostrand, Princeton, 1963.
- [50] Kenneth B. Stolarsky. Sums of distances between points on a sphere. *Proc. Amer. Math. Soc.*, 35:547–549, 1972.
- [51] M. E. Tytkin. On Hamming geometry of unitary cubes. *Soviet Physics. Dokl.*, 5:940–943, 1960.
- [52] Frederick A. Valentine. *Convex sets*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill Book Co., New York, 1964.

- [53] Stephen Willard. *General topology*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970.
- [54] D. J. Wilson. A game with squared distance as payoff. *Melbourne University Mathematics Research Report*, no. 16, 1982.
- [55] H. S. Witsenhausen. Metric inequalities and the zonoid problem. *Proc. Amer. Math. Soc.*, 40:517–520, 1973.
- [56] Reinhard Wolf. On the average distance property and certain energy integrals. *Ark. Mat.*, 35(2):387–400, 1997.
- [57] David Yost. Average distances in compact connected spaces. *Bull. Austral. Math. Soc.*, 26(3):331–342, 1982.